**INSIGHT**
*Frontiers in biology*

# nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

# PHILANTHROPY

## How to benefit from private funding—while avoiding the pitfalls

**PAGES 252, 254 & 260**

**ENVIRONMENT**

### THE AMAZON ON THE BRINK

*ll human action transform asin into a carbon source?*

**PAGES 239 & 321**

**AVIAN FLU**

### SENSE AND CENSORSHIP

*How should you publish data on dangerous lab-bred virus?*

**PAGE 257**

**COSMOLOGY**

### A DARK GALAXY UNVEILED

*Distant satellite supports cold dark matter theory*

**PAGES 271 & 341**

**➲ NATURE.COM/NATURE**

19 January 2012 £10

Vol. 481, No. 7381

03

9 770028 083095

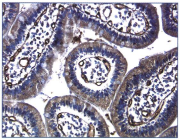
- Cancer
- Development
- Endocrinology
- Glycobiology
- Immunology
- Neuroscience
- Proteases
- Signal Transduction
- Stem Cells

**R&D Systems** Tools for Cell Biology Research™

# R&D Systems Specialized Tools for Cancer Research

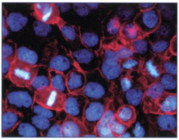
Cancer is a disease of cellular dysfunction involving a range of biological activities that promote unregulated proliferation. To help advance the understanding of this complex disease, R&D Systems offers an expansive selection of quality products, including **high performance antibodies, proteins, ELISAs, & arrays,** for investigating various cancer-related factors. Every product from R&D Systems undergoes rigorous quality control testing to ensure industry leading reliability.

### Angiogenesis



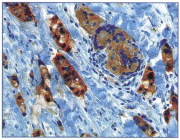
Angiopoietin-2 (Catalog # AF623)
Human Gastrointestinal Cancer Tissue

### Epithelial to Mesenchymal Transition



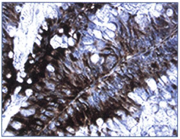
E-Cadherin (Catalog # AF648)
Human Epidermoid Carcinoma Cells

### Cancer Biomarkers



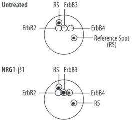
BRCA1 (Catalog # MAB22101)
Human Breast Cancer Tissue

### Inflammation



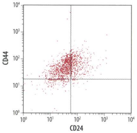
EMAP-II (Catalog # AF1910)
Human Colon Cancer Tissue

### Growth Factor/Receptor Signaling



Proteome Profiler™ *96* Phospho-RTK Antibody Array (Catalog # ARZ001)
MDA-MB-453 Cell Line

### Cancer Stem Cells



MagCellect $CD24^-CD44^+$ Breast Cancer Stem Cell Isolation Kit (Catalog # MAGH111)
MCF-7 Cell Line

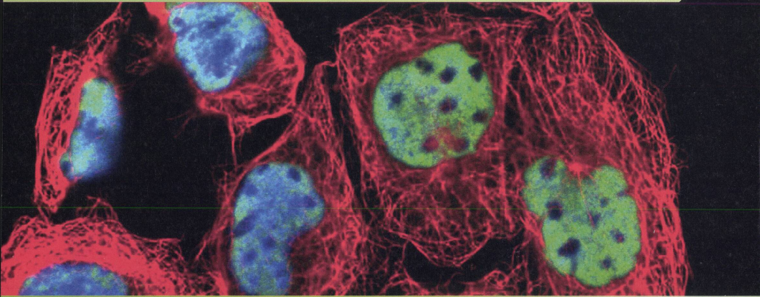
For more information visit our website at www.RnDSystems.com/go/Cancer

For research use only. Not for use in diagnostic procedures.

**R&D Systems, Inc.** www.RnDSystems.com
**R&D Systems Europe, Ltd.** www.RnDSystems.co.uk
**R&D Systems China Co., Ltd.** www.RnDSystemsChina.com.cn

**antibody**pedia | nature

Finding the right antibody for the right application just got easier!



*Antibodypedia* is a free resource that helps you find the right antibody for the right application. Independent, with data curated with the assistance of an international advisory board, *Antibodypedia* lets you:

- Search for antibodies that have proved themselves effective for specific applications
- Discover published research employing particular antibodies
- Submit antibody validation data from your own experiments

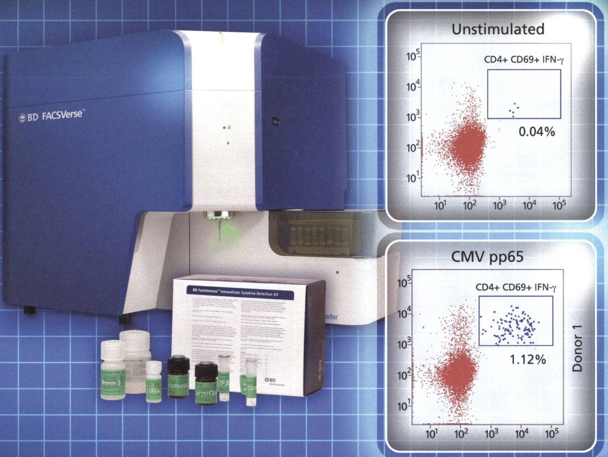
Find the right antibody for the right application by visiting:

**www.antibodypedia.com**

nature publishing group npg

# Detection of Intracellular Cytokines in T Lymphocytes on the BD FACSVerse™ System

## A streamlined approach to application support.



BD FACSVerse™
BD Fastimmune™ Intracellular Cytokine Detection Kit
Perm 2
EDTA
CD69/CD4
IFN-γ/CD
BFA
CD4
BD
Unstimulated
CD4+ CD69+ IFN-γ
0.04%
10^1
10^2
10^3
10^4
10^5
CMV pp65
CD4+ CD69+ IFN-γ
1.12%
Donor 1

**BD** Helping all people live healthy lives

## What's Really Exciting Is What's Behind it.

The predefined BD FastImmune™ 3-color CD4 intracellular detection kit combined with the BD FastImmune CD4 3C assay available for BD FACSuite software provides an easy and reliable way to detect cytokine-producing T lymphocytes upon antigen stimulation.

Using the assay, you can study baseline levels of CD69 expression and also the variation between donors in the CD4⁺CD69⁺IFN-γ⁺ response after stimulation with CMV pp65.

New BD FACSuite™ software makes it easy to convert experiments to reusable assays with setup information, acquisition and analysis templates, gating strategies, and reports you can share with other BD FACSVerse™ systems.

The software also provides BD predefined research assays matched with BD reagent kits for reproducible results. These assays reduce the variability of data between users and among laboratories for the same application.

Download a free application note and learn more about BD FACSuite predefined assays for apoptosis, cell cycle, cell proliferation, and cytokine detection at **bdbiosciences.com/go/verse.**

***Simply Brilliant.***

Class 1 Laser Product.
For Research Use Only. Not for use in diagnostic or therapeutic procedures.
BD, BD Logo and all other trademarks are property of Becton, Dickinson and Company. © 2011 BD
23-13665-00

**BD Biosciences**
2350 Qume Drive
San Jose, CA 95131
bdbiosciences.com

# CONTENTS

19 January 2012 / Vol 481 / Issue No 7381

## THIS WEEK

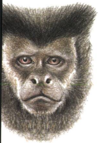
### EDITORIALS

237 **POLAR SCIENCE**
**Antarctic Treaty is cold comfort**
Scientists must take an active role to keep Antarctica for research

238 **MISCONDUCT**
**Face up to fraud**
Funders and universities need to find ways to tackle misconduct in the UK

### WORLD VIEW

239 **Break down boundaries in climate research**
*Paulo Artaxo*
Ambitious Amazon project points the way for collaboration over global warming



### RESEARCH HIGHLIGHTS

240 **SELECTIONS FROM THE SCIENTIFIC LITERATURE**
Lead-utilizing fungi / Why primates make faces/ Plants with a taste for roundworms / Aerosols linked to precipitation

### SEVEN DAYS

242 **THE NEWS IN BRIEF**
US stem-cell ruling challenged / Reprieve for San Raffaele / Grant odds low at NIH / Russian probe falls to Earth / Smallest tetrapod revealed

## CAREERS

399 **OUTREACH**
**Field hospitality**
Taking guests on field trips can be a rewarding experience — or not

401 **CAREER BRIEFS**

## NEWS IN FOCUS

245 **CLIMATE CHANGE**
Cutting smog offers fast way to mitigate global warming

247 **CLIMATE CHANGE**
EPA launches database of greenhouse-gas emitters

248 **EDUCATION**
Evolution advocate turns attention to climate change

249 **NUCLEAR POLITICS**
Assassinations unlikely to affect Iranian nuclear plans

250 **PHARMACEUTICALS**
Russian law has negative effect on drug approvals

251 **ASTRONOMY**
Reboot for Gemini Observatory puts it at the cutting edge

## PHILANTHROPY

PHILANTHROPY

### EDITORIAL

238 **Cap in hand**
The changing face of science funding

### NEWS FEATURES

252 **Like it? Pay for it**
Crowd-sourced funding is beginning to make waves

254 **Sponsor my science**
The art of marketing your research to prospective private sponsors

### COMMENT

260 **The price of charity**
*Patrick Aebischer*
The importance of sharing research overheads across all funding sources

## COMMENT

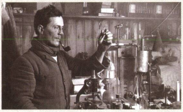
257 **BIOSECURITY**
**The fight over flu**
*Ron Fouchier & Ab Osterhaus; John Steinbruner; Kwok-Yung Yuen; D A Henderson; Lynn Klotz & Ed Sylvester; Jeffery K Taubenberger; Richard H Ebright & David L Heymann*
Expert opinion is divided on how to publish potentially dangerous data

### BOOKS & ARTS

**ANTARCTICA**

**Going to extremes**

Colin Martin reflects on the scientific legacy of Captain Scott's doomed expedition PAGE 264



261 **PUBLIC HEALTH**
**Biodefence built on teamwork**
*John Grabenstein*

262 **FICTION**
**Cosmic creation**
*Pedro Ferreira*

263 **BOOKS IN BRIEF**

### CORRESPONDENCE

265 Leap-year publicity for rare diseases / Many species threatened by appetite for traditional medicine / Managing fisheries / Drug patents sound

### OBITUARY

266 **Paul Mead Doty (1920–2011)**
*Stuart A Rice & Robert Haselkorn*

### FUTURES

404 **A game of self-deceit**
*Clayton Locke*

COVER IMAGE: 2&3 ILLUSTRATION

sysmex

Prospective Protein Provider

ProCube

Evolution comes, suddenly.

Bringing unprecedented efficiency to pharmaceutical development:
Introducing ProCube, a service that assists the analysis of protein interactions.

You want to outsource search and quality control processes. But you need the work done quickly, and you want to keep a lid on costs. ProCube is designed to meet the demands of this type of development environment. We offer high-quality, ready-to-use analytical tools based on proprietary specimen testing technology and protein production technology that uses silkworms. This approach streamlines business processes from search to quality control. The evolution of the pharmaceutical development environment starts with Sysmex.

**Features of ProCube**

● Validation data

We provide peace of mind and earn customers' trust by centralizing immobilization work and including interaction analysis results.

● Multi-variety, high-volume samples

We can accommodate a broad range of applications using silkworm technology.

procube.sysmex.co.jp

Sysmex Corporation

# CONTENTS

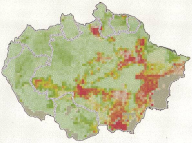
19 January 2012 / Vol 481 / Issue No 7381

## RESEARCH

CLIMATE SCIENCE

### All change

The Amazon basin is threatened by human activities. **PAGES 239 & 321**



## NEW ONLINE

267 Papers published this week at nature.com

## NEWS & VIEWS

268 MATERIALS SCIENCE
**A fresh twist for self-assembly**
Chirality controls colloid assembly of colloid particles
*Volker Schaller & Andreas R Bausch*
SEE LETTER P. 348

269 GENOMICS
**The path to retinoblastoma**
Epigenetic phenomena may underlie progression of childhood cancers
*Julien Sage & Michael L Cleary*
SEE ARTICLE P. 329

270 EVOLUTIONARY BIOLOGY
**A ratchet for protein complexity**
Evolution studied through reconstruction of ancestral proteins
*W Ford Doolittle* SEE LETTER P. 360

271 GALAXY FORMATION
**Distant dwarfs**
Discovery of a satellite galaxy at a cosmological distance
*Robert W Schmidt* SEE LETTER P. 341

273 CELL CYCLE
**A division duet**
A phosphatase and a kinase impose order during mitosis
*Curt Wittenberg*

274 ORGANIC CHEMISTRY
**Stabilizers cause instability**
Surprisingly reactive amides
*Jonathan Clayden*

275 CANCER
**Reprogramming clinical outcome**
Oestrogen receptor activity and breast tumours' fate
*Geneviève Deblois & Vincent Giguère*
SEE LETTER P. 389

## REVIEW

321 CLIMATE SCIENCE **The Amazon basin in transition**
*E A Davidson* et al.

## ARTICLES

329 GENOMICS **A novel retinoblastoma therapy from genomic and epigenetic analyses**
*J Zhang* et al. SEE N&V P.269

335 STRUCTURAL BIOLOGY **Structure of HDAC3 bound to co-repressor and inositol tetraphosphate**
*P J Watson, L Fairall, G M Santos & J W R Schwabe*

## LETTERS

341 ASTRONOMY **Gravitational detection of a low-mass dark satellite galaxy at cosmological distance**
*S Vegetti* et al. SEE N&V P.271

344 APPLIED PHYSICS **Coherent singlet-triplet oscillations in a silicon-based double quantum dot**
*B M Maune* et al.

348 APPLIED PHYSICS **Reconfigurable self-assembly through chiral control of interfacial tension**
*T Gibaud* et al. SEE N&V P.268

352 EARTH SCIENCE **Kimberlite ascent by assimilation-fuelled buoyancy**
*J K Russell, L A Porritt, Y Lavallée & D B Dingwell*

357 ECOLOGY **Recovery rates reflect distance to a tipping point in a living system**
*A J Veraart* et al.

360 EVOLUTION **Evolution of increased complexity in a molecular machine**
*G C Finnigan, V Hanson-Smith, T H Stevens & J W Thornton*
SEE N&V P.270

365 CELL BIOLOGY **Global landscape of HIV–human protein complexes**
*S Jäger* et al.

371 CELL BIOLOGY **Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection**
*S Jäger* et al.

376 VIROLOGY **T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction**
*W Zhang, J Du, S L Evans, Y Yu & X-F Yu*

380 BIOCHEMISTRY **Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia**
*C M Metallo* et al.

385 BIOCHEMISTRY **Reductive carboxylation supports growth in tumour cells with defective mitochondria**
*A R Mullen* et al.

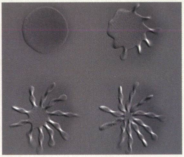
389 CANCER **Differential oestrogen receptor binding is associated with clinical outcome in breast cancer**
*C S Ross-Innes* et al. SEE N&V P.275

394 IMMUNOLOGY **Natural killer cells act as rheostats modulating antiviral T cells**
*S N Waggoner, M Cornberg, L K Selin & R M Welsh*

APPLIED PHYSICS

### Self build

A viral twist adds variety to colloid self-assembly. **PAGE 348**



CUSTOM MONOCLONAL ANTIBODIES
Rat mAb ◦ Mouse mAb ◦ Ascites

CUSTOM PROTEIN EXPRESSION
HEK293 ◦ Baculovirus ◦ *E. coli*

*Superior Quality • Unriveled Expertise*
**866.339.0871**
www.promab.com | info@promab.com

ProMab
Biotechnologies, Inc.

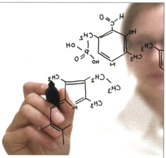
naturejobs

**Scientists stay informed of the latest:**
- **Hot jobs**
- **Career news**
- **Announcements**

Sign-up to our **FREE** jobseeker newsletter:
www.naturejobs.com/newsletter

MACMILLAN
SCIENTIFIC COMMUNICATIONS

Scientific ***Editing***



***Publishing help*** now available

MSC's Scientific Editing Service goes far beyond regular language editing and provides you with:

- in-depth **developmental editing** by ***Nature*-standard editors**
- **feedback** and suggestions on your manuscript's content, organisation and presentation by an expert in your field
- **advice on suitable journals** for your manuscript
- **language editing** by Nature Publishing Group Language Editing

**Submit** your manuscript today!

**www.mscediting.com**

'Nature Publishing Group editorial and publishing decisions are independent of MSC services.

*Exclusive partner of Nature Publishing Group, publisher of* Nature *and* Scientific American

SCIENTIFIC AMERICAN™

**Find out what's featured in every issue:**
Register now to receive your free monthly Table of Contents e-alerts

**www.nature.com/scientificamerican**

Gain insights into developments in science, technology and medicine

™ Scientific American, Inc.

nature publishing group npg

# INSIGHT

19 January 2012 / Vol 481 / Issue No 7381

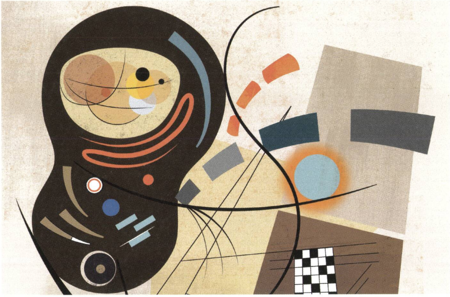
## FRONTIERS IN BIOLOGY

### REVIEWS

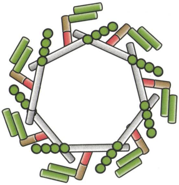
**278 Inflammasomes in health and disease**
Inflammation is the body's response to infection and tissue damage. Inflammasomes are a group of protein complexes that recognize inflammation-inducing stimuli (including pathogen-associated molecular patterns and damage-associated molecular patterns) and that control the production of pro-inflammatory cytokines. In addition, studies have revealed that inflammasomes have a diverse set of functions as part of the antimicrobial response, in metabolic syndrome and in inflammatory bowel disease. Mutations within inflammasomes are associated with a tendency to develop immune-mediated disease in humans. Although research has so far focused on *in vitro* and *in vivo* studies, future research is expected to centre on the contribution of altered inflammasome function to complex human disease.
*Till Strowig* et al.

**287 The DNA damage response and cancer therapy**
Genomic instability is a characteristic of most solid tumours and adult-onset leukaemia. When changes to the structure of DNA occur in the crucial driver genes within the tumour, these mutations can alter cell behaviour and drive the development of the disease. This instability is probably the result of several factors including DNA damage from environmental agents. The complex mechanisms that counteract threats to the genomic integrity are collectively known as the DNA damage response. The use of targeted therapies that exploit the DNA damage response could help to deliver a better therapeutic response in the future.
*Christopher J. Lord & Alan Ashworth*

**295 The promise of induced pluripotent stem cells in research and therapy**
Initial research suggested that the new, induced, pluripotent cells were functionally equivalent to embryonic stem cells; however, refined analysis has revealed that there are numerous molecular differences. The features of embryonic stem cells and induced pluripotent cells influence the use of each cell type in disease modelling and therapeutics. Understanding the genomic alterations that take place during the reprogramming of induced pluripotent cells is essential to ensure that the cells derived are functional, pure and appropriate for use in research and therapy.
*Daisy A. Robinton & George Q. Daley*



**306 Clonal evolution in cancer**
Cancer clone evolution takes place within tissue-ecosystem habitats, which have evolved to optimize multicellular function but restrain the clonal expansion of renegade cells. By introducing artificial selection in the form of drugs or radiation, oncologists change the cancer-clone dynamics and provide a selective pressure for the proliferation of variant cells that resist treatment. The tools of evolutionary biology can be used to explain the modest returns from cancer therapy and suggest alternative strategies for effective control, including focusing on the micro-environmental habitat and controlling cancer rather than attempting to eradicate it.
*Mel Greaves & Carlo C. Maley*



**The inflammasome complex.** PAGE 280

**314 The contribution of bone to whole-organism physiology**
Most organs have more functions than expected and this has led to the realization that a whole-organism approach to the study of physiology is required. Skeleton physiology is one example of how the function of one organ has a connection with other organs such as the brain, pancreas and gut. Leptin, serotonin, insulin and osteocalcin are all thought to be part of the pathways that coordinate the regulation of bone mass, energy metabolism and fertility. This connection between other organs and the regulation of bone mass is potentially important for the treatment of osteoporosis.
*Gérard Karsenty & Mathieu Ferron*

# nature

19 January 2012 / Vol 481 / Issue No. 7381

## EDITORIAL

**LONDON, MUNICH, WASHINGTON DC, NEW YORK, BOSTON, PARIS, SAN FRANCISCO, TOKYO** **nature@nature.com**

**EDITOR-IN-CHIEF:** Philip Campbell

**MANAGING EDITOR:** Nick Campbell

**PUBLISHING EXECUTIVE EDITOR:** Maxine Clarke **EXECUTIVE EDITOR:** Véronique Kiermer

**CHIEF MAGAZINE EDITOR:** Tim Appenzeller

**THIS WEEK/NEWS/FEATURES:** Alison Abbott, David Adam, Ananyo Bhattacharya, Geoff Brumfiel, Declan Butler, Ewen Callaway, Erika Check Hayden, Daniel Cressey, David Cyranoski, Natasha Gilbert, Eric Hand, Heidi Ledford, Corie Lok, Brendan Maher, Emma Marris, Richard Monastersky, Brian Owens, Helen Pearson, Mark Peplow, Eugenie Samuel Reich, Ivan Semeniuk, Quirin Schiermeier, Jeff Tollefson, Richard Van Noorden, Meredith Wadman, M Mitchell Waldrop **COMMENT:** Sara Abdulla, Joanne Baker, Rosalind Cotter, Barbara Kiser, Alison McCook, Lucy Odling-Smee, Sarah Tomlin **NATURE PODCAST/VIDEO:** Adam Rutherford, Eric Olson, Kerri Smith, Charlotte Stoddart **NEWS AND VIEWS:** Sadaf Shadan, Ana Lopes, Andrew Mitchinson, Cesar Sanchez, Marian Turner **RESEARCH:** Ritu Dhand, Karl Ziemelis, Francesca Cesari, I-han Chou, Tanguy Chouard, Rosamund Daw, Alex Eccleston, Angela Eggleston, Joshua Finkelstein, Henry Gee, Patrick Goymer, Noah Gray, Marie-Thérèse Heemels, Magdalena Helmer, Karen Howell, Claudia Lupp, Barbara Marte, Andrew Mitchinson, Juliane Mössinger, Deepa Nath, Leslie Sage, Magdalena Skipper, Clare Thomas, John VanDecar, Liesbeth Venema, Ursula Weiss, Michael White **INSIGHTS/REVIEWS/PERSPECTIVES:** Ursula Weiss **CAREERS:** Gene Russo, Karen Kaplan **TECHNOLOGY FEATURES:** Monya Baker **FUTURES:** Henry Gee **SUPPLEMENTS:** Herb Brody, Michelle Grayson, Tony Scully **SUBEDITORS:** Colin Sullivan, Sarah Archibald, Nicola Bailey, Anne Blewett, Nicolas Fanget, Isobel Flanagan, Paul Fletcher, Afsaneh Gray, Dinah Loon, Anna Novitzky, Yukie Ozawa, Katherine Perry, David Price, Jenny Rooke, Chris Simms, David Winters, Anna York **PRODUCTION:** Jenny Henderson, Leonora Dawson-Bowling, Dan Foley, Stewart Fraser, Susan Gray, Simon Gribbin, James McQuat, Emilia Orviss, Yvonne Strong, Robert Sullivan, Roberta Tarran, Charles Wenz **ART AND DESIGN:** Kelly Buckheit Krause, Wesley Fernandes, Madeline Hutchinson, Barbara Izdebska, Paul Jackman, Jasiek Krzysztofiak, Annthea Lewis, Nik Spencer, Claire Welsh, Mark Wood **JAPAN EDITION:** Mika Nakano, Akemi Tanaka, Mitsuyo Utsugi **PRESS OFFICE:** Ruth Francis, Neda Afsarmanesh, Lisa Boucher, Rachel Twinn, Rebecca Walton **ADMINISTRATION:** Paul Sng, Helen Anthony, Liz Bankes, Roseann Campbell, Ffion Cleverley, Faye Fornasier, Amy Invernizzi, Diane Kempinski, Hazel Mayhew, Katie McGoldrick, Jenny Meyer, Rina Nozawa, Kenneth Simpson, Claire Smith

## PUBLISHING

**LONDON** **feedback@nature.com**

**MANAGING DIRECTOR & PUBLISHER:** Steven Inchcoombe

**ASSOCIATE PUBLISHER:** Arash Hejazi

**PUBLISHING PROJECT MANAGER:** Christian Manco

**PROJECT MANAGER:** Helen Anthony

**TOKYO** **feedback@natureasia.com**

**PUBLISHING DIRECTOR — ASIA-PACIFIC:** David Swinbanks

**ASSOCIATE DIRECTOR — ASIA-PACIFIC:** Antoine E Bocquet

## DISPLAY ADVERTISING, SPONSORSHIP & NATUREJOBS

**GLOBAL HEAD:** Andrew Douglas a.douglas@nature.com

### DISPLAY ADVERTISING AND SPONSORSHIP

**GLOBAL HEAD:** Gerard Preston g.preston@nature.com

**ASSISTANT SPONSORSHIP MANAGER:** Reya Silao Tel: +44 (0)20 7843 4977 (r.silao@nature.com)

**EUROPE/REST OF WORLD display@nature.com Tel: +44 (0)20 7843 4960**

**GERMANY/SWITZERLAND/AUSTRIA:** Sabine Hugi-Fürst Tel: +41 52761 3386

**UK/IRELAND/SCANDINAVIA/SPAIN/PORTUGAL:** Evelina Rubio Hakansson Tel: +44 (0)20 7843 4079

**FRANCE/BELGIUM/THE NETHERLANDS/LUXEMBOURG/ITALY/OTHER EUROPE:** Nicola Wright Tel: +44 (0)20 7843 4959

**NORTH AMERICA** **display@natureny.com**

**NEW ENGLAND/ EASTERN CANADA:** Sheila Reardon Tel: +1 617 494 4900

**NEW YORK/MID-ATLANTIC/SOUTHEAST:** Jim Breault Tel: +1 212 726 9334

**MIDWEST:** Mike Rossi Tel: +1 212 726 9255

**WEST COAST/ WESTERN CANADA:** George Lui Tel: +1 415 781 3804

**WEST COAST NORTH:** Kayla McCutchan Tel: +1 415 403 9038

**ASIA-PACIFIC** **display@natureasia.com**

**ASIA-PACIFIC HEAD:** Kate Yoneyama

**JAPAN:** Ken Mikami Tel: +81 3 3267 8765

**GREATER CHINA/SINGAPORE:** Gloria To Tel: +852 2811 7191

**INDIA:** Vikas Chawla Tel: +91 124 288 1057

### SPONSORSHIP

**BUSINESS DEVELOPMENT MANAGER:** David Bagshaw Tel: +1 212 726 9215 **BUSINESS DEVELOPMENT MANAGER:** Patrick Murphy Tel: +1 617 475 9216

### NATUREJOBS

**naturejobs@nature.com**

**EUROPEAN HEAD OFFICE, LONDON:** Tel: +44 (0)20 7843 4961

**EUROPEAN SALES MANAGER:** Nils Moeller (n.moeller@nature.com) **BUSINESS DEVELOPMENT MANAGER, MIDDLE EAST:** Jon Giuliani (j.giuliani@nature.com) **ADVERTISING PRODUCTION MANAGER:** Stephen Russell **NATUREJOBS ONLINE PRODUCTION:** Dennis Chu

**US HEAD OFFICE, NEW YORK:** naturejobs@natureny.com **US SALES MANAGER:** Ken Finnegan

**JAPAN HEAD OFFICE, TOKYO:** Tel: +81 3 3267 8765 **ASIA-PACIFIC HEAD:** Kate Yoneyama

**ASIA-PACIFIC SALES & BUSINESS DEVELOPMENT MANAGER:** Yuki Fujiwara **BUSINESS DEVELOPMENT MANAGER, GREATER CHINA, SINGAPORE:** Gloria To Tel: +852 2811 7191

**INDIA:** Vikas Chawla Tel: +91 1242881057 (v.chawla@nature.com)

## MARKETING & SUBSCRIPTIONS

**USA/CANADA/LATIN AMERICA** **subscriptions@us.nature.com**

Tel: (USA/Canada)+1 866 363 7860; (outside USA/Canada) +1 212 726 9223

**MARKETING:** Sara Girard **FULFILMENT:** Karen Dawson

**JAPAN/CHINA/KOREA** **subscriptions@natureasia.com**

Tel: +81 3 3267 8751 **MARKETING/PRODUCTION:** Keiko Ikeda, Takesh Murakami

**EUROPE/REST OF WORLD** **subscriptions@nature.com**

Tel: +44 (0)1256 329242

**MARKETING:** Elena Woodstock, Marcus Parker, Hannah Phipps **FULFILMENT:** Jean Burrows

**INDIA** **npgindia@nature.com**

Tel: +91 124 2881053/54

**HEAD OF BUSINESS DEVELOPMENT, INDIA:** Debashish Brahmachari **MARKETING:** Harpal Singh Gill

**ANNUAL SUBSCRIPTIONS (UK/Europe, including post and packing)**

**FULL RATE:** £1,400

**PERSONAL RATES:** UK £135; Europe €209; Rest of World £210

**STUDENT RATES:** UK £79; Europe €122; Rest of World £125

Individual rates apply to payments by personal cheque or credit card. Back issues: £20.00.

**ANNUAL SUBSCRIPTIONS (USA, including post and packing)**

**INSTITUTIONAL/CORPORATE RATE:** $2,730 ; **PERSONAL RATE:** $199; **STUDENT RATE:** $99; **POSTDOC RATE:** $119

Printed in USA. Individual rates available only to subscribers paying by personal check or credit card. Orders for student/postdoc subscriptions must be accompanied by a copy of student ID. Rates apply to USA, Canada, Mexico/Central & South America. Add 7% GST tax in Canada (Canadian GST number 140911595). **BACK ISSUES:** US$20.00.

**ANNUAL SUBSCRIPTIONS (Asia, including post and packing)**

**JAPAN:** **年間購読料金：** *Nature のみ* 437,800 円 +21,890 円（消費税）=459,690 円（合計）
*Nature + Nature Digest* 466,550 円 +23,328 円（消費税）=489,878（合計）
お問い合わせ：日本出版貿易株式会社 国内営業１部
〒101-0064 東京都千代田区猿楽町1-2-1 Tel: 03-3292-3767 Fax: 03-3292-8766
**個人/学生購読について**
NPG ネイチャー アジア・パシフィック　フルフィルメント & カスタマーサービス
〒162-0843 東京都新宿区市谷田町2-37 千代田ビル Tel: 03-3267-8759 Fax: 03-3267-8752

**CHINA:** www.natureasia.com/ch/subscribe; **KOREA:** www.natureasia.com/korea/subscribe

## NATURE ONLINE

WWW.NATURE.COM/NATURE

**CHIEF TECHNOLOGY OFFICER:** Howard Ratner **ASSOCIATE DIRECTOR, NATURE.COM:** Daniel Pollock **PLATFORM DEVELOPMENT:** Jeremy Macdonald **WEB PRODUCTION:** Alexander Thurrell, Deborah Anthony **WEB DESIGN:** Juan-Carlos Sobrino, Luke Stavenhagen **PLATFORM INTERFACES:** Amanda Ward

## SITE LICENCES, FULFILMENT & CUSTOMER SERVICES

**SITE LICENCES:** npg.nature.com/libraries feedback@nature.com

**CHIEF OPERATIONS OFFICER:** John Carroll **ASSOCIATE DIRECTOR OF OPERATIONS:** Dominic Pettit

**CUSTOMER SERVICE:** Gerald Coppin **ORDER PROCESSING OPERATIONS:** Steve Poulter

## NATURE PUBLISHING GROUP OFFICES

**LONDON:** The Macmillan Building, 4 Crinan Street, London N1 9XW Tel: +44 (0)20 7833 4000 Fax: +44 (0)20 7843 4596/7

**BASINGSTOKE:** Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants RG21 6XS, UK Tel: +44 (0)1256 329242

**WASHINGTON DC:** 968 National Press Building, 529 14th St NW, Washington DC 20045-1938 Tel: +1 202 737 2355

**NEW YORK:** 75 Varick St, 9th Floor, New York, NY 10013-1917 Tel: +1 212 726 9200

**BOSTON:** 25 First Street, Suite 104, Cambridge, MA 02141 Tel: +1 617 475 9275

**SAN FRANCISCO:** 225 Bush Street, Suite 1453, San Francisco, CA 94104 Tel: +1 415 403 9027

**MUNICH:** Nymphenburger Strasse 14, D-80335 München Tel: +49 89 549057-13

**PARIS:** Moulin de Malicorne, Sainte Maure de Touraine, Paris 37800 Tel: +33 2 47 35 72 15

**TOKYO:** Chiyoda Building, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo 162-0843 Tel: +81 3 3267 8751

**INDIA:** Nature Publishing Group, 3A, 4th Floor, DLF Corporate Park, Gurgaon 122002 Tel: +91 124 2881053/54

*Nature*® (ISSN 0028-0836) is published by Nature Publishing Group, a division of Macmillan Publishers Ltd (The Macmillan Building, 4 Crinan Street, London N19XW). Registered as a newspaper at the British Post Office. North and South American orders to subscriptions@natureny.com (Tel: +1 866 363 7860). Other orders to subscriptions@nature.com (Tel. +44 (0)1256 329242; Fax +44 (0)1256 812358). Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the relevant copyright fee is paid direct to CCC, 222 Rosewood Drive, Danvers MA 01923, USA. Identification code for *Nature*: 0028-0836/03. CPC PUB AGREEMENT #40032744. In the US, *Nature* (ISSN 0028-0836) is published weekly on Thursday, except the last week in December by Nature Publishing Group, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. US Periodicals postage paid at New York NY, and additional mailing post offices. US POSTMASTER: send address changes to *Nature*, Subscription Dept, 75 Varick St, 9th Fl, New York NY 10013-1917, USA. Published in Japan by NPG Nature Asia-Pacific, Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843, Japan. © 2012 Macmillan Publishers Limited. All rights reserved.

# THIS WEEK

## EDITORIALS

**PHILANTHROPY** How to marry a millionaire with your science **p.238**

**WORLD VIEW** Giant Amazon study shows the future of climate research **p.239**

**SPECIES RECOGNITION** Social life leaves primates red-faced **p.240**

# Antarctic Treaty is cold comfort

*Researchers need to cement the bond between science and the South Pole if the region is to remain one of peace and collaboration.*

"All the day dreams must go; it will be a wearisome return." Written almost exactly a century ago, in the diary of doomed Antarctic explorer Robert Scott, those words mark the moment at which the British naval officer realized that the game was up. He had lost the race to the South Pole to his Norwegian rival Roald Amundsen, and the return journey was to prove worse than Scott imagined. By the end of March 1912, Scott and his two remaining companions had perished, just 18 kilometres from life-saving supplies. Their ill-fated expedition had a scientific slant too, and the zoologist Edward Wilson, who died alongside Scott, was the first scientist to see the South Pole.

It is perhaps a fitting legacy that the hostile landscape of Wilson's final months is now the scene of unprecedented scientific cooperation. Amid conflicting territorial disputes, the 1959 Antarctic Treaty handed the continent to the world's researchers, with the explicit goal of ensuring "in the interests of all mankind that Antarctica shall continue forever to be used exclusively for peaceful purposes and shall not become the scene or object of international discord". Some 30 countries now operate research bases in Antarctica, and the Iranian Students' News Agency reported last week that Iran intends to open one within three years.

The international rush to Antarctica in the name of science has not gone unnoticed by some with interests beyond research. In an article published in *The Australian* newspaper on 31 December, Sam Bateman and Anthony Bergin of the Australian Strategic Policy Institute in Canberra raised the provocative prospect that countries such as China and India could use bases recently opened there to improve satellite communications to military forces that increasingly depend on space-based infrastructure. "To do so would be at odds with the Antarctic Treaty, but the sparse use of the treaty's inspection mechanisms means that such activity could go undetected," Bateman and Bergin say. "If Antarctic sites take on military significance, we could see a move towards destabilisation of Antarctica as a zone of peace."

That may seem a far-fetched scenario, but events may not have to take such a dramatic turn to undermine the treaty and weaken the monopoly that science currently has on Antarctica. The late Christopher Joyner of the Department of Government at Georgetown University in Washington DC identified three potential challenges to the cooperative spirit of the agreement, driven by the changing global political and economic climate. States might implement national continental-shelf claims in offshore Antarctic waters in pursuit of energy resources, he suggested, or tensions could escalate between Japan and Australia over whaling in Antarctic waters. Joyner's third scenario — widespread and unregulated bioprospecting — is a topic already identified as problematic by treaty members, and one discussed at their annual meeting last summer, held in Buenos Aires. Nearly 200 research organizations from 27 states are carrying out research for commercial purposes in the Antarctic, Joyner said, and one big goal is sifting its plants and animals for beneficial genetic and biochemical resources.

**➲ NATURE.COM**
How exploring launched Antarctic science:
**go.nature.com/ppfdmu**

Although the Antarctic Treaty seems under no immediate threat, one need only look at the escalating political importance of the Arctic to see the potential for change in the south. Just this week, members of the Arctic Council were due to discuss requests from China and India for official observer status. And even under the strict conditions of the treaty, the appeal of Antarctica to some nations searching for new mineral and energy resources has never been too far from the surface.

**"There remains great appeal in the spirit of Scott's expedition — of scientists as explorers."**

Scientists can play an important part in preserving the treaty and protecting their unique Antarctic playground, and the promise it continues to offer for research. Interest in the poles is currently sky-high — witness the success and international syndication of the BBC's *Frozen Planet* series — and the more that scientists can promote the work that they do there, the more the bond between the two will be cemented in the public's mind. There remains great appeal in the spirit of Scott's expedition, of scientists as explorers, and in that vein, sometime this year, researchers in Antarctica are likely to be the first to penetrate a sub-glacial lake. The long-standing Russian effort to drill into Lake Vostok stalled again this month, raising the prospect that, if they fail to break through this month, then a British team seeking to investigate Lake Ellsworth could beat them to it in December. Not that priority — the loss of which so devastated Scott — matters in 2012. "It's not a race," a spokeswoman for the British project insists. We believe them, just. ■

# Face up to fraud

*The UK government and funding agencies must address research misconduct.*

Many people in science would rather not talk about the problem of research misconduct, much less act on it. After all, who directly involved would benefit from a serious crackdown? Certainly not the institutions at which the misconduct takes place — they are nominally responsible, but can face legal repercussions, embarrassing headlines and a public-relations disaster if they expose cheating academics. It is much easier to shuffle miscreants out of the side door with vague references and a promise of silence, effectively pushing the problem somewhere else, and onto someone else.

So it is perhaps a sort of progress that the *British Medical Journal* and the international Committee on Publication Ethics were able to organize a meeting on the subject in London last week, gathering representatives from universities, funders, journals and lobby groups to discuss how the problem could be tackled in the United Kingdom (see *Nature* http://doi.org/hmx; 2012). The meeting broke little new ground, but its organizers do, at least, deserve credit for trying.

A big part of the problem is the lack of perceived risk associated with misconduct. Some fraudulent researchers might be sociopaths who don't care about the rules, but many others simply believe that they can anticipate the outcome of a research project, and see no downside to fabricating the required results to save time, or tweaking results to achieve a stronger signal. Either way, stronger action and punishments are needed to discourage such misbehaviour. (Meanwhile, for colleagues considering blowing the whistle, the risks are glaringly huge — witness the plight of scientists, such as cardiologist Peter Wilmshurst, who have raised questions and have faced the full force of Britain's ludicrous libel laws as a result.)

Could publications such as this one do more to deter cheats? Unfortunately, we are often in no position to flag up even proven cases of misconduct, and thereby highlight the risks that miscreants run with their careers. Yes, it is a journal's primary job to clean up the literature, but when papers are retracted owing to misconduct, the libel laws (again) often prevent our editors from saying so. We know that this leaves the affected communities frustrated and in the dark. It leaves us frustrated, too.

So, with journals unable to push towards greater integrity and universities often unwilling to do so, should funding agencies be leading the charge? It is, after all, their money that is wasted if misconduct does occur.

Funding agencies in the United States do sometimes investigate misconduct. Research funded by the National Institutes of Health and some other government agencies falls under the remit of the Office for Research Integrity (ORI), which has the power to bar researchers from receiving future funding. However, as Nicholas Steneck, director of the research-ethics programme at the Michigan Institute for Clinical and Health Research in Ann Arbor, told the London meeting, this process probably misses most major misconduct. And the ORI can't initiate investigations: institutions must conduct their own inquiries first.

> **"Stronger action and punishments are needed to discourage misbehaviour."**

In the United Kingdom, there seems to be little appetite for launching an overarching ORI-type regulator. Certainly, the existing independent advisory group, the UK Research Integrity Office in Falmer, is clear that it has no desire to take on such a role. British funding councils — in collaboration with the country's universities — have chosen instead to produce a 'concordat' detailing good practice, to which institutions will be expected to sign up. This is laudable, but unlikely to strike fear into fraudsters and fabricators.

So, how can Britain highlight cases of misconduct and discourage it in future? Ultimately, the incentives probably need to come from on high, and the government could get the ball rolling by commissioning an anonymous survey on misconduct that UK researchers have witnessed and perpetrated. An official audit would offer a strong platform for others to build on — perhaps with a parliamentary inquiry and subsequent report on the damage done to UK science by misconduct, and an assessment of the options for tackling it and the investment needed. Funders and universities could then work together to establish common definitions of what counts as misconduct, and how it will be punished. And if a reform of the libel laws goes ahead, journals and other scientists would be able to do more to highlight and expose miscreants.

Sounds ambitious? If the solutions were easy, there wouldn't be a problem to discuss. But there is, so we must face it. ■

---

# Cap in hand

*A word to the wise on getting that much-needed research funding.*

Would you be willing to spend weekends on the yacht of a friendly billionaire in the name of science? Or insist to airport check-in staff that your life-saving research demands that you be upgraded from economy to business class? Perhaps you would be happy to see your face on a T-shirt? Or for folks you met on the Internet to traipse through your lab, taking photographs?

Welcome to the cold reality of science in a global recession. As the flow of public money slows across the world, academic researchers are increasingly turning to private funders and wealthy individuals-turned-philanthropists to pay for their work. There is a strong tradition of such support already, of course, especially in the biomedical field. But as government cuts around the world begin to bite, more and more scientists will be looking for alternative sources of income.

So, in a short series of articles this week, *Nature* focuses on where that money is and how you can access it. Do not feel too proud to ask. Some of your competitors are doing it already.

In our News Features, we look at the two extremes of research paid for by private individuals — from the billionaires willing to set up entire laboratories and pay for the work done there for years, to the web-based begging bowls that can take just a few dollars each from thousands of different people.

On page 254, we talk to those at the top: the big spenders, the entrepreneurs — and those scientists who have benefited from their largesse. How did they do it? Partly by being in the right place at the right time, although it helps to know where the right places are. And it helps even more to have something to say when you get there. As Thomas Pierson of the SETI Institute in Mountain View, California, says, "People give money to people." One secret seems to be to think big. "If I ask for $100,000 and they say 'yes' right away, then I didn't ask for enough," one university fund-raiser tells us. "It's a common mistake."

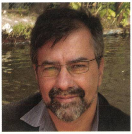
At the other end of the scale are the crowd-funding websites, explored on page 252, where scientists can post details of their projects and ask many individuals to collectively cough up for them. From a low-cost robot for tackling oil spills to a project to map water quality along the Mississippi River, some researchers are already adept at tapping the potential of the masses. And although the target donors may be different, one key strategy remains the same: tell a good story. Sell yourself and sell your science.

Still, be wary of selling yourself short, warns a Comment piece on page 260. Patrick Aebischer, president of the École Polytechnique Fédérale de Lausanne in Switzerland, complains that too many donations to university research from charities and foundations come with a catch — they don't pay for the associated costs, such as salaries and utility bills. "Institutions with many privately funded projects are effectively 'punished' for their success," he writes. "To meet the higher research-infrastructure costs, universities may drain resources from education, or diminish 'expensive' disciplines such as physics, chemistry or engineering, in which philanthropic support is scarce." The solution, he says, is for institutions to identify the full cost of research activities and pass it on. "Private bodies should not hijack university resources. They should contribute a fair share of the expense."

The money is certainly out there. Just look at the billions poured into football teams. And if the pages of the glossy magazines can be believed, the luxury-goods market remains strong. The money must be spent on something, so why not science? And although those weekends on luxury yachts may be a tough way to make it happen, someone has to do it. ■

**➲ NATURE.COM**
To comment online, click on Editorials at:
go.nature.com/xhunqv

WORLD VIEW *A personal take on events*


A. P. FREIRE

# Break down boundaries in climate research

*Scientists wanting to implement change must collaborate between disciplines. An ambitious Amazon study shows them how, says* **Paulo Artaxo**.

Tropical forests are big players in the global balance of climate and carbon, and the Amazon is the biggest of the lot. Almost 20 years ago, I was part of a group of scientists that hatched an ambitious plan to study the Amazon rainforest in a totally new way, to improve our understanding of the forest and its role in the global climate system. We wanted to build a scientific base from which to learn how the region could be developed sustainably.

I say it was ambitious, but perhaps that is an understatement. Our plan was to study the rivers and soil, use satellites to track changes in land use, monitor the atmosphere with aircraft and analyse the socio-economic factors that drive deforestation. And we wanted to install more than a dozen enormous monitoring towers deep in the remotest parts of the jungle, to measure changes in the levels of trace gases, aerosols, carbon and water vapour, among other things, and to study energy balance and plant metabolism.

At that time, environmental science in Brazil was in its infancy. But we were very optimistic, as most Brazilians are, with no reason to be sure it would work.

Today, the success of the Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA), as the project is known, shows that we were right to be optimistic. So, too, does the catalogue of scientific papers and results that our experiment produced over the years, which are detailed in a Review on page 321 of this issue. In fact, the LBA is a triumph for long-term cooperative research. Its success has led to calls for similar projects to be set up to monitor other tropical forests, including those in Africa and southeast Asia. So, here is how we did it.

RESEARCH ON CLIMATE CHANGE CAN **BE EFFECTIVE** ONLY IF IT IS ACCOMPANIED BY **GREATER INTEGRATION.**

Perhaps most fundamentally, the project encouraged cooperation between physicists, chemists, meteorologists and biologists, to tease apart the Amazon's interwoven scientific strands. The biology of the forest, for instance, drives emissions of volatile organic compounds, which affect cloud formation and development, changing how much sunlight reaches Earth and thus affecting the radiation balance. Unusually for the time, we also brought economists and social scientists into the research. We knew that this was necessary if we were to fully understand all the factors influencing such a complex system.

International cooperation has also been consistently required: each project team includes scientists and students from several countries, and every few years, the LBA holds its own scientific conference in Brazil, which attracts hundreds of scientists and students. For a single experiment, such large participation is outstanding.

The LBA was a pioneer in data sharing and access. From the beginning, policies encouraged our researchers to make information widely available, and to invite other scientists to use LBA data in their models. This helped to boost our output to more than 2,000 papers in two decades. The project has also produced 200 PhDs and nearly 300 master's degrees, and many graduates have continued to work in the Amazon region. As a result, several courses on environmental physics, ecology, climate and environment have been launched at universities in the area.

➲ NATURE.COM
Discuss this article online at:
go.nature.com/sldalu

The LBA was launched with domestic and international funds (especially from NASA), and Brazilian funding agencies strongly supported scholarships to help students to participate. The project has survived several changes of government in Brazil. LBA research is now funded by the Ministry of Science and Technology and administered by the National Institute of Amazonian Research in Manaus, aided by growing awareness of regional and global environmental issues and sustainable development. In a fast-changing scientific and political field, the LBA and its leaders have been flexible. As the climate questions have become more important, we have seen that we are in a unique position to answer important questions, such as how aerosol particles emitted by burning biomass could affect cloud microphysics and precipitation, as well as how the increasing risks of fire from changes in land use and global warming interact to alter forest ecosystems. Expansion of soya farming into the Amazon was not significant when the LBA began, but when it later became so, our biophysical and socio-economic researchers altered their projects to study it.

Twenty years after it was conceived, the LBA has provided the scientific knowledge and personnel we need to help to manage the Amazon. It has provided evidence that the environmental and economic value of standing forest is significantly higher than that of degraded areas. And the large decrease in annual deforestation in Amazonia, from more than 27,000 square kilometres in 2004 to 6,200 square kilometres in 2011, shows that it is possible to implement strategies to reduce deforestation and carbon emissions in the region.

Can the LBA experience be exported to other regions? It needs to be. Research on tropical ecosystems and their role in climate change can be effective only if it is accompanied by greater integration of natural and social sciences. We need to break down the boundaries between university departments and funding agencies that cater only to broad disciplines, such as 'sciences' or 'arts'. Integrated science at the international level is certainly the way to go. This is urgent: the planet calls for action now. ■

**Paulo Artaxo** *is professor of environmental physics at the University of São Paulo in Brazil.*
*e-mail: artaxo@if.usp.br*

# RESEARCH HIGHLIGHTS

*Selections from the scientific literature*

**MICROBIOLOGY**

## Bacterial weapon: inflammation

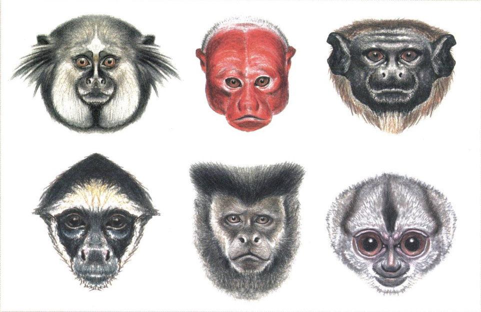
Inflammation is normally bad news for pathogens, but one bacterium seems to have turned the process to its own advantage.

Alan Hauser and his team at Northwestern University in Chicago, Illinois, studied a particularly virulent strain of the pneumonia-causing *Pseudomonas aeruginosa* that had been isolated from a patient. The bacterium produces a protein known as RhsT, and the authors determined that *P. aeruginosa* injects this protein into host cells called phagocytes — white blood cells that normally engulf and destroy pathogens. Once inside them, RhsT activates an immune-signalling complex called the inflammasome, killing infected cells.

The authors found that the RhsT protein is linked to increased production of inflammatory proteins in cell culture and in the lungs of infected mice. The mice died within two days of infection. Conversely, mice infected with *P. aeruginosa* in which the *rhsT* gene had been deleted survived.

Many bacterial pathogens contain genes similar to *rhsT*, which may be important for infection.

***Proc. Natl Acad. Sci. USA*** **http://dx.doi.org/10.1073/pnas.1109285109 (2012)**



S. NASH

**EVOLUTION**

## Social life shapes primate faces

Facial patterns such as skin and hair colour may have evolved to help primates recognize and communicate with others of their species.

Sharlene Santana and her colleagues at the University of California, Los Angeles, analysed skin and hair colour patterns in the faces of adult males of 129 species of primate from the Americas (a sampling pictured). The authors found that species living in smaller groups have more complex facial features. This suggests that these animals, which interact with each other less frequently than those living in larger groups, are more dependent on facial patterns to recognize members of their own species. Living near a greater number of related species also contributes to facial diversity, regardless of group size.

However, ecological factors also play a part — species range, for example, seems to drive the evolution of hair length and the pigmentation of certain facial features.

***Proc. R. Soc. B*** **http://dx.doi.org/10.1098/rspb.2011.2326 (2012)**

**PHYSICS**

## Not so ice-free materials

Surfaces engineered to repel water and ice can lose their ice resistance if environmental conditions shift.

'Superhydrophobic' surfaces may prevent ice build-up and have applications in aerospace and power transmission, among others. Dimos Poulikakos at the Swiss Federal Institute of Technology in Zurich and his colleagues ran water droplets down four different hydrophobic surfaces at −15 °C, in an experimental chamber in which humidity and nitrogen gas flow rate could be controlled. Although the droplets did not freeze at the water–surface boundary, at certain levels of humidity and gas flow they did freeze at the water–gas boundary. This occurred because of evaporation at this interface that further cooled the water, and resulted in much higher adhesion of ice to the surface under particular conditions.

***Nature Commun.*** **http://dx.doi.org/10.1038/ncomms1630 (2012)**

**BIOCHEMISTRY**

## Fungi munch on lead

Two species of fungus can transform metallic lead, a toxic pollutant, into a stable lead mineral called chloropyromorphite. The formation of this mineral has been proposed as a way to sequester lead contaminants in soil.

Geoffrey Gadd at the University of Dundee, UK, and his colleagues incubated small spheres of lead with either *Metarhizium anisopliae* or *Paecilomyces javanicus*, fungal species isolated from a former lead-mining area. They found mineral deposits on the surface of the spheres, and X-ray analysis of these minerals revealed signatures of chloropyromorphite after one month of incubation. The mineral became the dominant transformation product after three months. Lead not exposed to fungi showed typical signs of corrosion.

This is the first time that microbes have been found to generate pyromorphite.
***Curr. Biol.*** **http://dx.doi.org/10.1016/j.cub.2011.12.017 (2012)**

**BOTANY**

## Plant gobbles buried worms

The Brazilian plant *Philcoxia minensis* feasts on nematode worms — making this spindly plant the first known to trap its prey using sticky leaves buried underground.

*Philcoxia* (**pictured**) grows in the white, nutrient-poor sands of the Brazilian Cerrado, a biodiversity hot spot. Rafael Oliveira at the State University of Campinas in São Paulo, Brazil, and his colleagues found dead nematodes trapped on *P. minensis*'s subterranean leaves. The team fed the plants nematodes containing labelled nitrogen. After two days, about 15% of the prey's labelled nitrogen was found in the leaves.

This, as well as the presence of enzymes called phosphatases on the leaves, suggests that the plant actively digests its prey.
***Proc. Natl Acad. Sci. USA*** **http://dx.doi.org/10.1073/pnas.1114199109 (2012)**

R. S. OLIVEIRA

**MICROBIOLOGY**

## How parasites get sticky

To invade host cells, one group of parasites relies on a protein that promotes the release of sticky molecules from a specialized organelle called a microneme.

Parasites of the phylum Apicomplexa, which includes those responsible for malaria (*Plasmodium falciparum*) and toxoplasmosis (*Toxoplasma gondii*), secrete a mixture of adhesion proteins at their surfaces, allowing them to stick to host cells. Marc-Jan Gubbels at Boston College in Massachusetts and his colleagues sequenced the genome of a mutant strain of *Toxoplasma* with an impaired ability to release its microneme contents or invade cells. They found a mutation in a protein called DOC2.1. Shutting off the gene encoding an equivalent protein in *Plasmodium* also impaired this parasite's ability to invade.
***Science*** 335, **218–221 (2012)**

**ATMOSPHERIC SCIENCE**

## More aerosols, more rain

Atmospheric aerosols seem to boost the intensity of rainfall over large areas of land and ocean.

Aerosols — tiny particles that include black carbon and sulphates — influence cloud properties, but whether the overall effect produces more or less rain has been uncertain. Ilan Koren at the Weizmann Institute of Science in Rehovot, Israel, and his colleagues looked at satellite observations of rainfall, aerosols and clouds between June and August 2007, and compared them with meteorological data. They found that increases in aerosol abundance correlated with a higher rate of rainfall in the tropics, subtropics and mid-latitudes.

The authors say that the similar trends seen across different locations and environmental conditions suggest a link between increased aerosol levels and more intense rainfall.
***Nature Geosci.*** **http://dx.doi.org/10.1038/ngeo1364 (2012)**

## COMMUNITY CHOICE

*The most viewed papers in science*

**DEVELOPMENTAL BIOLOGY**

## Watching cells die in real time

**✪ HIGHLY READ** on jcb.rupress.org 11 Dec–10 Jan

A type of programmed cell death called apoptosis has been imaged in live mouse embryos, providing insight into a developmental process that can go awry and cause neural birth defects.

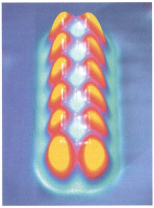
Yoshifumi Yamaguchi and Masayuki Miura at the University of Tokyo and their team engineered mice that generate a fluorescent signal in response to the activation of caspases, enzymes that trigger apoptosis. The authors used a fast-scanning confocal microscope to detect this activation in developing mouse embryos at single-cell resolution. They also conducted time-lapse imaging of an early-developmental process called neural-tube closure — in which apoptosis occurs and the basic structure of the central nervous system is formed.

In embryos that were missing key apoptotic genes, this closure process occurred much more slowly and, in some embryos, even failed to complete. The authors suggest that neural-tube closure must occur within a specific time window to be completed properly.
***J. Cell Biol.*** 195, **1047–1060 (2011)**

**APPLIED PHYSICS**

## Magnetic switch for memory

Some data-storage devices rely on ferromagnets — in which the neighbouring magnetic spins of atoms are aligned — to hold information as magnetic bits. However, magnetic fields can corrupt stored data. Antiferromagnets, in which the neighbouring spins are oppositely aligned, are not as sensitive to magnetic fields, so are a promising alternative. Sebastian Loth, Andreas Heinrich and their colleagues at IBM's Almaden Research Center in San Jose, California, have read and switched the magnetic states of antiferromagnetic iron atoms, demonstrating that information can be stored antiferromagnetically.



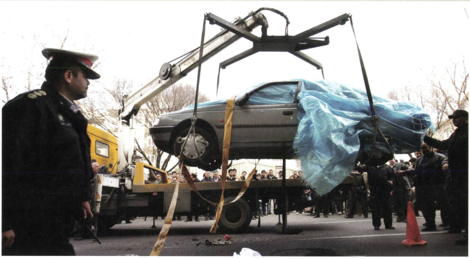
S. LOTH, IBM RESEARCH/ALMADEN

The authors placed a few antiferromagnetic iron atoms in a regular pattern (**pictured**) on a copper nitride surface, and switched their magnetic states by applying a current of more than 7 millivolts to individual atoms using the tip of a scanning tunnelling microscope. From images taken with the microscope, the authors could distinguish two stable and distinct magnetic states of the atoms.
***Science*** 335, **196–199 (2012)**

**➲ NATURE.COM**
For the latest research published by *Nature* visit:
**www.nature.com/latestresearch**

# SEVEN DAYS

*The news in brief*



MEGHDAD MADADI, FARS NEWS AGENCY/AP

## Iranian chemical engineer assassinated

A chemical engineer who worked at Iran's Natanz uranium-enrichment facility was killed on 11 January in Tehran — the latest victim of a string of assassinations and attacks apparently aimed at Iran's nuclear programme. Mostafa Ahmadi Roshan Behdast, 32, was assassinated by a magnetic bomb that had been attached to his car (pictured). He was identified in the Iranian press as a deputy director of marketing at the Natanz facility. See page 249 for more.

### EVENTS

## Mars probe crash

Russia's Phobos-Grunt spacecraft, which failed to make it out of Earth orbit in its attempt to reach a moon of Mars last year, splashed down in the southern Pacific Ocean on 15 January. As *Nature* went to press, exact details of the probe's crash and location were unclear, but it is expected to have scattered fragments over thousands of square kilometres. On-board tanks containing unused, toxic fuel probably broke up well before the probe re-entered Earth's stratosphere. See go.nature.com/qdv3to for more.

### POLICY

## NOAA to move?

US President Barack Obama has proposed moving the National Oceanic and Atmospheric Administration (NOAA) from the Department of Commerce to the Department of the Interior, as part of a broader agency reorganization that is intended to save taxpayers about US$3 billion over 10 years. Obama made the proposal on 13 January; what the move would mean for NOAA is unclear. See go.nature.com/efbprv for more.

## Stem-cell lawsuit

Two scientists seeking to block US government funding for research using human embryonic stem (ES) cells filed their arguments to a US appeals court on 12 January. James Sherley and Theresa Deisher, who work on adult stem cells, are appealing a July 2011 decision in which a federal judge ruled against their case. The appeal is set to be heard on 23 April. On 12 January, the US National Institutes of Health made four more human ES-cell lines eligible for use in federally funded research, bringing the approved total to 146. See go.nature.com/wtsnxn for more.

## US reactor safety

The United States may follow France in recommending that nuclear power plants install back-up equipment for containing a serious accident, in the wake of the meltdowns at Japan's Fukushima plant. On 11 January, the Nuclear Regulatory Commission (NRC) said that it considered a plan put forward by the Nuclear Energy Institute, a body based in Washington DC that represents the nuclear industry, to be "a reasonable starting point". The plan involves placing portable pumps, generators and other equipment around reactors for use in emergencies. The NRC is expected to announce new safety measures before the 11 March anniversary of the Fukushima accident. See go.nature.com/saeouu for more.

## Emissions data

The US Environmental Protection Agency has released a public database of greenhouse-gas emissions from the nation's largest power plants and industrial facilities. The agency requires more than 6,200 facilities to report annual emissions, and their 2010 emissions cover around half of the nation's direct emissions. See page 247 for more.

### FUNDING

## San Raffaele saved

The debt-ridden San Raffaele Scientific Institute in Milan, one of Italy's most prestigious biomedical research institutes, will not be taken over by the Vatican after all. The institute filed for bankruptcy last October, after accumulating debts of €1.5 billion (US$1.9 billion), and a consortium led by the Vatican Bank put in a bid to rescue it (see *Nature* **478,** 296–297; 2011). But Milanese businessman Giuseppe Rotelli, head of the private-health group San Donato, which runs 18 hospitals in northern Italy, unexpectedly outbid the consortium on 10 January. Scientists had been concerned that Vatican influence could have distorted research agendas.

## Low odds at the NIH

The success rate for scientists seeking grants from the US National Institutes of Health in Bethesda, Maryland, fell to a historic low of 18% in 2011, the agency announced on 13 January. The decline was driven by an 8% increase since

2010 in applications submitted for competitive awards, to almost 50,000. At 15%, women's success rates were equal to those of men for new grant applications, but when competing for the renewal of existing grants, women's 33% success rate trailed men's 36%.

## Clean-energy spend

Global government spending on research and development in clean energy fell by 22% from 2010 to US$13.2 billion last year, according to figures released by analysts Bloomberg New Energy Finance on 12 January. Corporate research spending also slipped back, falling by 14% from 2010 to $13.2 billion. But total investment in the sector — which includes the financing of energy projects such as large solar installations and rooftop photovoltaics — surged by 5% from 2010 to reach $260 billion worldwide.

**RESEARCH**

## Frogs in size war

Actual size
Previous record

C. AUSTIN/LOUISIANA STATE UNIV.

Barely a month after a frog took the title of world's tiniest tetrapod, an even smaller frog has emerged to steal the crown. With an average length of just 7.7 millimetres from snout to vent, *Paedophryne amauensis* (**pictured**) is 1 mm shorter than the previous record-holder, *P. dekot*. Both frogs live in leaf litter in the forests of Papua New Guinea and make up two of the six minuscule *Paedophryne* species so far identified on the island. The latest find was reported on 11 January (E. N. Rittmeyer *et al. PLoS ONE* http://doi.org/hmw; 2012).

## Transgenic flight

The German chemical giant BASF is shifting its transgenic-plant operations from Europe to the Americas, it says, because of widespread opposition to the technology from consumers, farmers and politicians at home. On 16 January, the company said that it would move its plant-science headquarters from Limburgerhof, Germany, to Raleigh, North Carolina, cutting 140 jobs in Europe, although it would retain research centres in Berlin and Ghent, Belgium. Jonathan Jones, a plant researcher at the Sainsbury Laboratory in Norwich, UK, said the move was "a sign that Europe is not open for business in this area". See go.nature.com/nramri for more.

## Nuclear security

Australia is the most secure of the 32 countries that hold at least one kilogram of weapons-useable nuclear material, according to a first-of-its-kind study that asked experts to assess risks of nuclear theft. North Korea is listed last, and the United Kingdom ranks highest among nuclear-armed states, in tenth place. The 11 January study (www.ntiindex.org) from the Nuclear Threat Initiative, a non-profit organization in Washington DC, also examined security conditions in 144 non-nuclear states that might be used as safe havens or transit points for stolen material.

**PEOPLE**

## Astrophysicist dies

Steven Rawlings, an astrophysicist at the University of Oxford, UK, who helped to instigate and coordinate the proposed Square Kilometre Array (SKA) telescope project, died in mysterious circumstances on 11 January, aged 50. His body was found at the house of a friend and police are now investigating the death. Shocked scientists added that it was a great loss to the SKA project. Telescope construction would not start until 2016, but a pivotal decision about whether to locate it in South Africa or Australia is expected by March. See go.nature.com/c8bhhn for more.

## COMING UP

**22–27 JAN**
The sixth 'Arctic Frontiers' conference, in Tromsø, Norway, discusses the potential of energy resources in the Arctic.
**www.arcticfrontiers.com**

**23 JAN–17 FEB**
The World Radiocommunication Conference of the International Telecommunication Union in Geneva, Switzerland, will vote on whether to abolish the leap second, pulling our reference time out of sync with the Sun (see *Nature* **479,** 158; 2011).
**go.nature.com/qq1zi7**

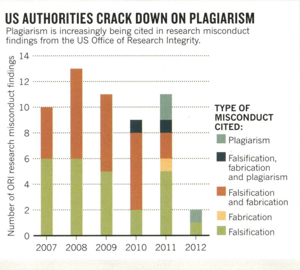
## Red-wine fraud

A three-year investigation into a biology laboratory at the University of Connecticut has found its chief guilty of falsifying and fabricating data in at least 23 papers and 3 grant applications. Dipak Das, director of the Cardiovascular Research Center at the University of Connecticut Health Center (UCHC) in Farmington, is a cardiologist whose studies included work on the red-wine chemical resveratrol. On 11 January, the UCHC issued a 60,000-page report detailing extensive misconduct; on the same day, Das was dismissed as co-editor-in-chief of *Antioxidants & Redox Signaling*, and two of his articles in the journal were retracted. See go.nature.com/faxbyj for more.

**➲ NATURE.COM**
**For daily news updates see:**
**www.nature.com/news**

## TREND WATCH

The US Office of Research Integrity (ORI) is sharpening its eye for plagiarism. In the past, the office has censured researchers mostly for fabricating (making up) and falsifying (manipulating or omitting) data, but two of its three latest findings involved plagiarism. John Dahlberg, head of investigative oversight at the ORI, says that such cases are surfacing because of increased use of plagiarism-detection software, so the office is toughening its stance. See go.nature.com/qvp8qg for more.


**US AUTHORITIES CRACK DOWN ON PLAGIARISM**
Plagiarism is increasingly being cited in research misconduct findings from the US Office of Research Integrity.

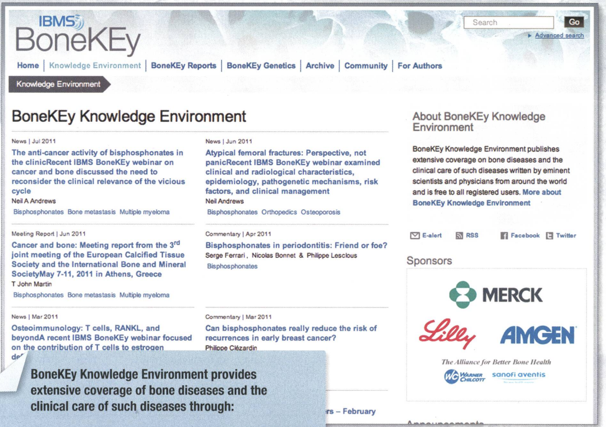
SOURCE: ORI

IBMS BoneKEy

# BoneKEy Knowledge Environment

## The premier online global network for bone health professionals

GET THE TAG READER FOR YOUR PHONE: **GETTAG.MOBI**
SCAN & REGISTER. **FREE ACCESS TO REGISTERED USERS.**



IBMS BoneKEy

Search Go

Advanced search

Home | Knowledge Environment | BoneKEy Reports | BoneKEy Genetics | Archive | Community | For Authors

Knowledge Environment

BoneKEy Knowledge Environment

News | Jul 2011
**The anti-cancer activity of bisphosphonates in the clinicRecent IBMS BoneKEy webinar on cancer and bone discussed the need to reconsider the clinical relevance of the vicious cycle**
Neil A Andrews
Bisphosphonates Bone metastasis Multiple myeloma

News | Jun 2011
**Atypical femoral fractures: Perspective, not panicRecent IBMS BoneKEy webinar examined clinical and radiological characteristics, epidemiology, pathogenetic mechanisms, risk factors, and clinical management**
Neil Andrews
Bisphosphonates Orthopedics Osteoporosis

Meeting Report | Jun 2011
**Cancer and bone: Meeting report from the 3rd joint meeting of the European Calcified Tissue Society and the International Bone and Mineral SocietyMay 7-11, 2011 in Athens, Greece**
T John Martin
Bisphosphonates Bone metastasis Multiple myeloma

Commentary | Apr 2011
**Bisphosphonates in periodontitis: Friend or foe?**
Serge Ferrari, Nicolas Bonnet & Philippe Lesclous
Bisphosphonates

News | Mar 2011
**Osteoimmunology: T cells, RANKL, and beyondA recent IBMS BoneKEy webinar focused on the contribution of T cells to estrogen de**

Commentary | Mar 2011
**Can bisphosphonates really reduce the risk of recurrences in early breast cancer?**
Philippe Clézardin

rs – February

About BoneKEy Knowledge Environment

BoneKEy Knowledge Environment publishes extensive coverage on bone diseases and the clinical care of such diseases written by eminent scientists and physicians from around the world and is free to all registered users. **More about BoneKEy Knowledge Environment**

E-alert RSS Facebook Twitter

Sponsors

MERCK
Lilly AMGEN
The Alliance for Better Bone Health
Warner Chilcott sanofi aventis

**BoneKEy Knowledge Environment provides extensive coverage of bone diseases and the clinical care of such diseases through:**

- High-quality opinion and analysis
- News & Commentaries
- Meeting Reports
- Webinars
- Web Portals: BoneKEy Genetics and BoneKEy Oncology (forthcoming in 2012)
- Multimedia article templates
- BoneKEy Community

Adhering to the highest standards of publication in the scientific community, all content on the *IBMS BoneKEy* platform is subject to rigorous peer review, copyediting and production procedures.

**www.nature.com/bonekey**

International Bone
& Mineral Society
IBMS

nature publishing group npg

# NEWS IN FOCUS

**EDUCATION** Veteran of evolution wars will fight for climate science **p.248**

**IRAN** Assassinations seen as unlikely to slow nuclear efforts **p.249**

**ASTRONOMY** New director reboots a lagging observatory **p.251**

**FUNDING** Pitching research to the crowd — and asking it to pay **p.252**

J. STANMEYER/VII/CORBIS

Soot emissions from sources such as this brick kiln in Karkhla, Pakistan, absorb sunlight and contribute directly to global warming.

GLOBAL WARMING

# Pollutants key to climate fix

*Governments look to reduce methane and black carbon as a way to slow warming.*

**BY JEFF TOLLEFSON**

Buses spew clouds of black exhaust fumes in Mexico City while, in India, wood burnt in rudimentary stoves fills houses with sooty smoke. Methane leaks from gas pipelines in Russia and rice paddies in China, eventually breaking down in sunlight and contributing to the production of smog and ozone. In each of these cases, simple steps to curb air pollution would promote public health; scaled up, they may offer the only realistic way to tame global warming over the next few decades.

Rapid measures to reduce emissions of black carbon, which soaks up solar energy, and methane, a greenhouse gas that is 25 times more potent than carbon dioxide, could cut the rate of global warming in half between now and 2050, according to an analysis published last week (D. Shindell *et al. Science* **335,** 183–189; 2012). Such numbers have spurred political interest, and next month a small coalition of countries is aiming to launch an initiative that would target these 'short-lived climate forcers'. If successful, the effort could have an immediate impact on global temperatures while countries grapple with efforts to regulate emissions of carbon dioxide, the most important greenhouse gas.

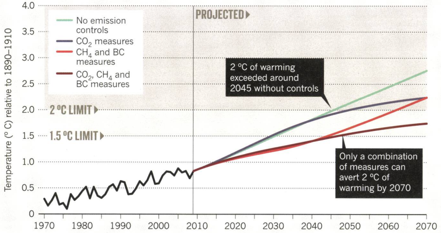
➲ **NATURE.COM**
For more on black carbon and its effects, visit.
go.nature.com/9khxtj

"We're in a gridlock over carbon dioxide, and we're losing time," says Veerabhadran Ramanathan, an atmospheric scientist at the Scripps Institution of Oceanography in La Jolla, California, and a co-author on the *Science* analysis. "This is one way to buy back some of that time, and the co-benefits are huge." By 2030, these reduction measures could prevent anywhere from 700,000 to 4.7 million premature deaths from air pollution annually, the study found. And because ozone is toxic to plants, such measures could boost global crop production by 1–4%.

The United Nations Environment ▶

## RAPID RESPONSE

Measures to reduce emissions of black carbon (BC) and methane ($CH_4$) would have an immediate effect on atmospheric warming. Controls on carbon dioxide ($CO_2$) are still needed to rein in temperature in the long run.



SOURCE: D. SHINDELL *ET AL. SCIENCE* **335,** 183–189 (2012)

▶ Programme explored the potential gains in a detailed assessment last June (see go.nature.com/4wcwxf). Chaired by Drew Shindell of the NASA Goddard Institute for Space Studies in New York, the assessment ranked hundreds of options for reducing black carbon and ozone pollution according to their potential to reduce warming. A follow-up report, released in November and funded by the Swedish government, further analysed opportunities and impacts at national and regional levels. This work served as the basis for the *Science* study.

For methane, the study identified 14 control measures that would target leakage from coal mining and oil and gas operations, emissions from landfills, wastewater systems, livestock manure and rice paddies. Black-carbon reduction would focus on cleaning up diesel vehicle emissions, biomass stoves, brick kilns and coke ovens. Other measures would reduce the burning of agricultural waste and provide alternatives to wood, dung and charcoal for cooking and heating in poor countries.

It could take decades to slow global warming through reductions in carbon dioxide emissions, whereas cutting soot and methane would have immediate climate payoffs because they are quickly purged from the atmosphere. Modelling suggests that, together, a 75% reduction in black carbon and a 40% reduction in methane emissions over the next two decades would lower projected warming by about 0.5 °C by 2050. However, aggressive mitigation of carbon dioxide would still be needed to keep the atmosphere from warming more than 2 °C, the current target of international climate negotiations (see 'Rapid response').

The assessment shows that the benefits of such reductions far outweigh the costs, with premature deaths from pollution predicted to fall the most in India and China. Cuts to black carbon would yield the biggest reduction in warming in Africa, south Asia and the Arctic, where the temperature rise would be reduced by two-thirds during the next three decades.

> ***"This is one of the most important climate agendas for the world right now."***

Kevin Trenberth, a climate scientist at the National Center for Atmospheric Research in Boulder, Colorado, credits the study with building scientific consensus on the issue, but warns that it should not detract from the main agenda on carbon dioxide. "The fundamental problem with long-term climate change is $CO_2$, and anything that takes us away from addressing that doesn't really solve the problem," he says. "It just puts it off."

Advocates concede as much but argue that both tracks are necessary. They point out that the agenda for short-term forcers can move forward now without a global climate treaty, which remains a distant prospect after last month's climate talks in Durban, South Africa, delayed setting new targets for emissions reductions (see *Nature* **480,** 299–300; 2011). Bilateral or multilateral agreements can help to pay for cleaner fuels and filters for diesel engines or promote policies that ban the burning of agricultural waste and promote aeration of rice paddies to decrease methane production. "It doesn't require 150 nations to sign a document," says Ramanathan. "It requires regional partnerships."

In fact, countries exploring methane and black-carbon mitigation are intentionally keeping the agenda separate from the UN climate negotiations. Following initial meetings last year hosted by the United States, Mexico and Bangladesh, these countries, as well as Sweden, Canada and Ghana, are hoping to work out some of the details of an initiative to promote mitigation steps at a diplomatic meeting in Montreal, Canada, next week. The initiative could involve a funding mechanism, a clearing house for projects or even a dedicated organization. "Sweden, Canada and the United States can be expected to put funds into the platform initially," says Caroline Dickson, who heads climate issues at the Swedish environment ministry. "Once we have launched this initiative other countries will be able to join." The goal is to unveil the initiative in Washington DC next month, Dickson says.

The initiative may seek to build on existing efforts. In October 2010, a group of 41 governments and international agencies created the Global Methane Initiative to help poor countries finance profitable methane-capture projects at landfills and other sites. Similarly, last year, various countries and companies launched the Global Alliance for Clean Cookstoves to help improve the quality of stoves in the developing world.

Although the measures under consideration are technically feasible today, aggressive implementation will require high-level political support and broader consideration at international forums such as the G8 and G20 talks, says Durwood Zaelke, who heads the Institute for Governance and Sustainable Development in Washington DC. "My worry is that they are going to undersell it," he says. "This is one of the most important climate agendas for the world right now, but it is not fully appreciated at the highest political levels what it can deliver." ■

## MORE ONLINE

**TOP STORY**

Totally drug-resistant TB emerges in India **go.nature.com/vrvgyk**

ATUL LOKE/PANOS

**MORE NEWS**

- Multicellular yeast show speed of evolution **go.nature.com/3grkip**
- British science needs 'integrity overhaul' **go.nature.com/duj1uo**
- Tests to extract methane from gas hydrates begin in Alaska **go.nature.com/mobk76**

**PODCAST**

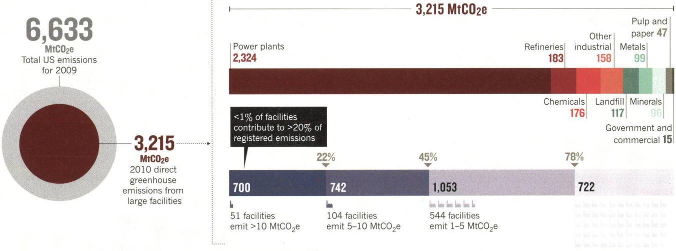
nature podcast

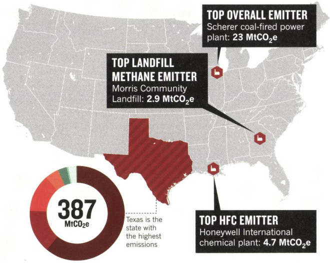
How intelligence ages, science in the Amazon, and crowd-sourcing your research funding **go.nature.com/p5vigz**

SOURCE: EPA

## THE GAS TRACKER

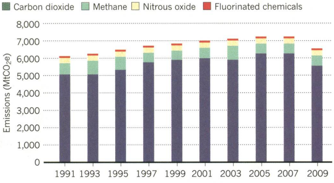
The US Environmental Protection Agency's online greenhouse-gas database covers about half of the country's 2010 direct emissions. About three-quarters of those are produced by power plants; around 45% come from 155 facilities (just 2.5% of those that submitted data), each of which emits at least 5 million tonnes of carbon dioxide equivalent (MtCO$_2$e) a year.




TOP OVERALL EMITTER
Scherer coal-fired power plant: 23 MtCO2e
TOP LANDFILL METHANE EMITTER
Morris Community Landfill: 2.9 MtCO2e
TOP HFC EMITTER
Honeywell International chemical plant: 4.7 MtCO2e
387
MtCO2e
Texas is the state with the highest emissions

## RISE AND FALL

Greenhouse-gas emissions in the United States have waned slightly owing to a downturn in the economy, as well as a move from coal to natural gas for power generation.



Summed subtotals differ from national total because of rounding.

**CLIMATE CHANGE**

# Database tallies US emissions

*Environment agency launches searchable public log of major greenhouse-gas emitters.*

**BY JEFF TOLLEFSON**

Keeping track of a country's greenhouse gases is an accounting problem of epic proportions. In the United States, scientists have relied on a mix of methods to build up their national emissions inventory, including monitoring the electricity output of a power plant and assessing the quality of the fuel that powers it.

Now they have a new resource: official data from the companies themselves, collated into a user-friendly online database that was launched on 11 January by the US Environmental Protection Agency (EPA).

"It's a great resource, and I'm sure people will find interesting things to do with it," says Gregg Marland, a geologist at Appalachian State University in Boone, North Carolina, who led the development of guidelines for the Intergovernmental Panel on Climate Change on how to construct and use national emissions inventories.

The inventory covers industrial, commercial and government facilities that emit more than 25,000 tonnes of carbon dioxide equivalent per year, and was prompted by a law passed in 2008 that was meant to support the growing efforts towards regulating greenhouse gases. Plans for that regulatory programme collapsed in 2010, but EPA officials continued to work on a database, which would still provide investors and consumers with statistics to help them pressure industry to cut emissions.

The inventory covers facilities that directly emit about half of the country's total emissions, but does not include emissions from agriculture or land use. The online database allows users to compare and rank about 6,200 facilities by state, sector, and type and volume of greenhouse gases, including the ubiquitous ▶

▶ carbon dioxide and methane and the more esoteric fluorinated chemicals.

The data should help researchers trying to measure and track greenhouse-gas emissions, as well as those studying natural carbon cycles. As scientists pin down exactly where anthropogenic carbon emissions are coming from, Marland says, they can separate out carbon uptake and emissions by plants and soils. "The better we understand both the magnitude and distribution of human emissions, the better we understand what is happening in the biosphere."

### THE USUAL CULPRITS

Some database searches produce very familiar results. Carbon dioxide makes up about 95% of the greenhouse-gas emissions logged (although if emissions from agriculture and other sectors in the full US inventory are included, that proportion would drop to around 83%). The state with the largest overall emissions is Texas, which logged 387 million tonnes of carbon dioxide equivalent in 2010. And power plants overshadow any other stationary sources of greenhouse gases, accounting for about three-quarters of emissions in the inventory (see 'The gas tracker').

But there are also some eye-opening statistics. Just 2.5% of the facilities that have submitted data to the EPA are responsible for 45% of the emissions, for example. And almost all of the United States' emissions of trifluoromethane (also known as HFC-23), a potent greenhouse gas, come from just two facilities — the Honeywell International plant in Baton Rouge, Louisiana, and the Dupont manufacturing plant in Louisville, Kentucky. Since 1990, US producers have voluntarily reduced by about 85% their emissions of trifluoromethane, which is a by-product of the manufacture of the refrigerant and chemical feedstock chlorodifluoromethane.

The flood of data does not mean that scientists can stop measuring greenhouse-gas emissions in the atmosphere. Pieter Tans and his team at the National Oceanic and Atmospheric Administration's Earth System Research Laboratory in Boulder, Colorado, measure greenhouse-gas plumes from major facilities through a network of tall monitoring towers, for example. The annual totals in the database will certainly help to improve their atmospheric models, he says, but monitoring how emissions vary over the course of hours and days is still vital.

Totalling those real-time measures should also provide a way to verify the companies' annual estimates. "The hypothesis is that these are correct emissions estimates," Tans says. "We can test that to see whether what is being reported is consistent with actual observations." ■

---

EDUCATION

# Evolution advocate turns to climate

*Education centre known for battling creationists aims to help science teachers convey understanding of global warming.*

BY SUSAN YOUNG

Students walking out of classrooms when global warming is mentioned; teachers pressured to change lesson plans to avoid the subject or portray it as speculative rather than a matter of scientific consensus. For Eugenie Scott, the stories and anecdotes fit a familiar pattern. Scott is executive director of the National Center for Science Education (NCSE) an organization based in Oakland, California, with a reputation for doggedly defending the teaching of evolution in US classrooms. But a growing impression that climate science is facing a similar struggle, together with entreaties from educators and textbook authors, has helped to convince her that the NCSE should expand its mandate to include the politically charged issue of global warming.

"I think we can make an important contribution," says Scott. "If teachers understand that there is a place that they can go to for help, we can use some of the expertise that we've gained over the years dealing with evolution to apply to this related problem."

A recent survey in the United States suggests that there is indeed a problem. From August to October 2011, the nation's National Earth Science Teachers Association (NESTA) queried Earth- and space-science teachers about their experiences of climate-change education. Depending on the region, 25–30% of respondents reported that students, parents, administrators or other community members had argued with them that climate change is not happening or that it is not the result of human activity. Some school boards and state legislators have threatened to require educators to 'teach the controversy' about climate change — a term coined in relation to evolution that amounts to presenting a scientific theory as one of various possible viewpoints.

Many of the teachers surveyed have strong science backgrounds and professional experience related to climate change. But "when you look at what the teachers are facing across the country, it goes way beyond science", says Roberta Johnson, executive director of NESTA, based in Boulder, Colorado. "It goes into areas of political debate."

In a 16 January announcement, the NCSE says that it will offer support to educators facing ideological opposition when teaching climate change, providing advice on how to present the underlying science. The strategy mirrors its approach to evolution, which includes clarifying for students why science is an appropriate tool for understanding the natural world. "This perspective is also important in helping people to understand the reasons why scientists overwhelmingly accept climate change," the NCSE says in a mission statement describing the new effort.

The statement also says that the NCSE will not take a position on what, if anything, should be done to counteract global warming or mitigate its effects. "What to do about it ranges widely and gets outside of the strict science and into policy issues in which many, many variables are going to have to be considered," says Scott. "We are not a policy think tank; we don't have expertise in this area."

But such policy neutrality may not prevent science teachers from being challenged in the classroom. "The core issue is not whether global warming is happening, or whether humans are involved, but whether it is a crisis," says James Taylor, a research fellow at the Heartland Institute, a libertarian think tank based in Chicago, Illinois, that opposes the regulation of carbon emissions (see *Nature* **475,** 440–441; 2011).

> **"When you look at what the teachers are facing across the country, it goes way beyond science."**

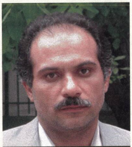
Scott acknowledges that there is more to teaching climate change than explaining the science clearly. "We need to be aware of the fact that people are very emotionally concerned about these issues," she says. If people feel threatened ideologically, politically or economically, "all the science in the world won't convince them". She adds that the NCSE will also help teachers to understand the views of parents and others who oppose the teaching of climate change.

"Knowing the motivations behind a parent's views can help a teacher come up with a solution or response that might assuage that parent's concerns and let their kid remain in the classroom," says Scott. ■

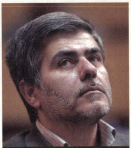
L TO R: FARS NEWS/REUTERS; B. MEHRI/AFP/GETTY; A. KENARE/AFP/GETTY; AP; FARS NEWS/AP

**NUCLEAR FALLOUT**

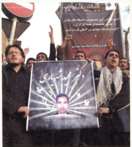
Several scientists have been targeted for assassination in Iran over the past two years, although some have no verifiable connection to the country's nuclear programme.



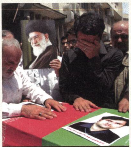
Masoud Alimohammadi
Particle physicist at the University of Tehran, no link to nuclear programme.
Killed 12 January 2010



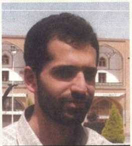
Fereydoun Abbasi-Davani
Senior nuclear scientist at defence ministry, now head of nuclear programme.
Injured 29 November 2010



Majid Shahriari
Nuclear physicist, possibly involved in nuclear programme.
Killed 29 November 2010



Dariush Rezaei-Nejad
Electrical-engineering student, no link to nuclear programme.
Killed 23 July 2011



Mostafa Ahmadi Roshan Behdast
Deputy director of marketing at Natanz uranium facility.
Killed 11 January 2012

NUCLEAR POLITICS

# Murders unlikely to slow Iran's nuclear efforts

*Experts say international sanctions are the best way to stall the weapons programme.*

BY SHARON WEINBERGER IN BEIRUT

Last week's assassination of an official working at Iran's uranium-enrichment facility was worthy of a James Bond film: the 32-year-old was killed by a magnetic bomb placed on his car by a passing motorcyclist.

But the murder of Mostafa Ahmadi Roshan Behdast is unlikely to result in a neat filmic denouement. His death is the latest in a string of assassinations and other attacks seemingly aimed at Iran's nuclear programme over the past few years (see 'Nuclear fallout').

Although experts agree that at least some of the killings are part of an organized foreign campaign to slow Iran's efforts to enrich uranium, they are sceptical that the strategy will work. "The immediate effect is very small," says Olli Heinonen, a senior fellow at Harvard University's Belfer Center for Science and International Affairs in Cambridge, Massachusetts. "If I have a project that is important for national security, I never count on one single person," he says, suggesting that international sanctions are a more effective way of slowing Iran's efforts.

Heinonen, former deputy director-general of the International Atomic Energy Agency in Vienna, points out that there is still no proof of who ordered the assassination, or why. Ahmadi Roshan was named in many Western press reports as a nuclear scientist (his degree was actually in chemical engineering), but the Iranian press identified him as deputy director of marketing at the uranium-enrichment facility in Natanz, suggesting that he was not a crucial figure in Iran's nuclear programme.

Peter Zimmerman, a physicist and emeritus professor at King's College London, says that killing key personnel could delay a nuclear programme in the right circumstances. For example, killing Leslie Groves, the military chief of the Manhattan Project during the Second World War, might have had serious consequences if it had happened in 1942 or 1943 when the atomic-bomb project was in doubt and Groves' advocacy was crucial. But "if I had killed [J. Robert] Oppenheimer in 1944, I don't think it would have delayed Hiroshima by even a month, and maybe not a week", suggests Zimmerman. In Iran, he says, none of the victims so far is crucial to the nuclear programme.

Even identifying key personnel in the programme can be difficult, because scientists suspected of involvement often have academic affiliations. In an interview in 2007 with this reporter, Fereydoun Abbasi-Davani introduced himself as a university professor and head of the Nuclear Society of Iran, an academic association for promoting nuclear knowledge. But he was also on a United Nations sanctions list for his alleged involvement in Iran's clandestine military programme. In 2010, an attempt to assassinate Abbasi-Davani failed, and he was subsequently appointed as the head of Iran's Atomic Energy Organization in Tehran.

At the time, Abbasi-Davani described efforts to slow Iran's nuclear programme as "fruitless", and resulting in only short-term setbacks. For example, he said, the United States allegedly led a ban on selling software to Iran that could be used for calculating nuclear reactions. Abbasi-Davani said that Iranian scientists eventually developed their own software, in part using information from contacts abroad. "We were slowed down, but we could get the knowledge and know-how that we needed in other ways," he said.

A similar recovery was seen after the Stuxnet computer worm wreaked havoc with the uranium-enrichment centrifuges at Natanz in 2010, says Scott Kemp, an associate research scholar at Princeton University's Woodrow Wilson School for Public and International Affairs in New Jersey. "They overcompensated, they began enriching even faster, and within six weeks, they were caught up," he says. Soon, he adds, the facility had moved even closer to being able to produce enough highly enriched uranium for a nuclear weapon.

Kemp suggests that killing scientists who are allegedly associated with the programme could have the same effect, by bolstering Iran's resolve to develop its nuclear capability. "These things have unintended and unforeseeable consequences," he notes. ■

PHARMACEUTICALS

# Russian drug law hinders clinical trials

*Legislation to increase availability of new medicines has delayed approvals.*

**BY ALLA KATSNELSON**

A law aimed at improving Russians' access to innovative medicines has backfired badly.

'On the circulation of medicines', a law that came into effect in September 2010, was intended to streamline the bureaucratic system for testing and registering new drugs in Russia. There were high hopes that it would also boost the nascent domestic pharmaceutical industry and attract foreign companies to run trials there. The law set a maximum cost and time limit for drug registrations, for example; mandated compensation for those injured during trials; and declared that drugs manufactured in Russia solely for export did not need to undergo registration.

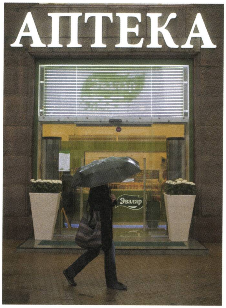
But the law has actually hampered clinical trials and limited the number of drug approvals over the past year, according to a report from the Association of Clinical Trials Organizations (ACTO) in Moscow, which represents companies conducting clinical trials in Russia.

One problem is that medicines must now undergo local trials before they can be sold in Russia, regardless of whether they have already undergone trials or been approved elsewhere. The goal is to ensure that drugs are backed by high-quality trials that prove their effectiveness in the indigenous population. A few Asian countries have similar rules, but whereas some Asians are thought to metabolize certain drugs in a different way from Europeans, that is not the case for Russians, says Leonid Kokovin, a clinical-trials manager based in Moscow. "We are statistically like European and American subjects, so the argument for testing on our population isn't understandable," he says.

The law also raises ethical concerns because it subjects people to clinical trials that are not expected to yield extra knowledge, Kokovin notes. He suggests that medicines backed by suitable foreign trials should be approved by the Russian authorities, as long as they are tracked in post-marketing studies. "That would be much simpler, and wouldn't put such a barrier on the market," he says.

ACTO's report describes one (anonymized) European company's struggle to register a drug for preventing preterm delivery, a condition that affects about 5% of pregnancies. The medicine is already approved for this use in many Western countries, and for other uses in Russia. The cost of running a Russian trial was estimated to be €1 million (US$1.3 million), and patients in the control group would have to be asked to avoid the drug, raising their risk of preterm delivery. Furthermore, clinical trials on pregnant women are rare in Russia, so there are few accredited facilities that can screen participants.

Despite the difficulties, the company is going ahead with the trial. However, many others have been discouraged by a lack of detail about how to comply with the law. "No one can say — including, probably the regulatory agency — what kinds of trial specifically should be run," says Dmitry Margolin, co-chairman of ACTO's regulatory committee. "No one can give the guarantee that if you do these trials then the data will allow you to register the medicine." Margolin says he is aware of "several tens" of instances in which manufacturers had planned to seek registration fora medicine, but had been deterred by the requirement for fresh trials.



**Russian pharmacies can't sell some internationally approved drugs, owing to a strict law.**

S. KRASILNIKOV/ITAR-TASS

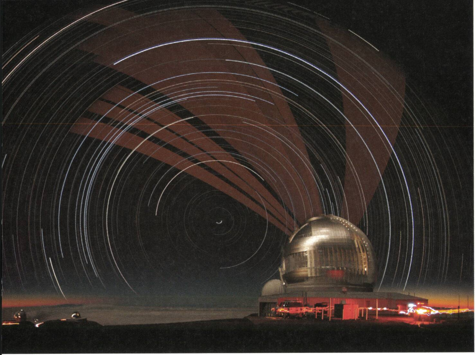
According to information from ACTO and the Ministry of Health and Social Development, which oversees the registration of medicines, just three repeat trials are under way for innovative medicines already approved outside Russia — two for hepatitis C and one for neuralgia — with two more set to start soon. "If the situation continues," says Margolin, "then a large number of new products will not be registered and will not reach Russian pharmacies."

On the basis of meeting summaries posted on the ministry website, ACTO also estimates that the ministry's Ethics Council rejects more than 30% of applications to run trials, mostly as a result of administrative or clerical errors in the applications. To avoid undue influence, Russia's new medicines law forbids the Ethics Council from directly contacting companies, so it cannot request that small errors are corrected. By contrast, in Germany, where the rejection rate is 1%, approvals are routinely conditional on small changes in the application. "It's one of the main flaws of the Russian regulatory system under the new law," says Sergei Lomakin, a senior associate at the Moscow office of the international law firm Baker & McKenzie. "The system leads to a bigger number of rejections than one that allows scientific dialogue" between companies and regulators, he says.

Another provision of the law — that medicines developed abroad cannot have their first clinical tests in Russia — is hampering foreign companies that are trying to develop products specifically for the Russian market. Ironically, many such companies have agreements with government-backed investors, such as the nanotechnology venture fund Rusnano, to develop new medicines in the country.

"These changes in the guidelines happened at the moment when the country had invested so much in innovative drug development that there were quite a few drugs which were around the corner from starting phase I trials," says Andrei Gudkov, chief scientific officer of Cleveland BioLabs in Buffalo, New York, which partners with Rusnano and is struggling to launch a Russian early-stage trial for one of its anticancer compounds. The Ministry of Health and Social Development did not to respond to *Nature*'s questions on the matter.

ACTO and others say that scrapping the requirement for local registration trials would be an important first step towards fixing the situation. Until that can be done, says Margolin, "companies for the most part are sitting and waiting". ■



Gemini South fires a laser at the sky (shown in time-lapse) to correct for atmospheric distortions.

J. POLLARD/GEMINI OBSERVATORY

ASTRONOMY

# Gemini's twin telescopes reboot

*Adaptive optics put the observatory at the cutting edge.*

**BY ERIC HAND**

The Gemini Observatory's star has been slow to rise. The publicly funded observatory boasts two of the world's largest telescopes, and offers a rare opportunity to watch both the northern and southern skies from sites in Hawaii and in Chile. Yet it has lagged in terms of scientific productivity since it opened in 2000, and millions of dollars have been spent on overambitious, highly specialized instruments that were ultimately cancelled. "It felt like a child who had gone astray," says Frederic Chaffee, who witnessed Gemini's struggles while director of the rival Keck Observatory on Hawaii's Mauna Kea. "It was frustrating to watch."

But less than a year after Chaffee was drafted out of retirement to become Gemini's interim director when his predecessor stepped down, the once faltering facility may finally be living up to its promise. Last week, at a meeting of the American Astronomical Society in Austin, Texas, Chaffee showed off results from several new instruments. These should benefit a range of observers at the twin 8.1-metre telescopes.

"In the last year, Gemini has made huge progress in turning around the observatory," says David Silva, director of the US National Optical Astronomy Observatory in Tucson, Arizona.

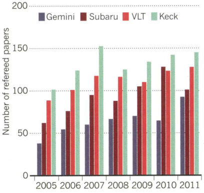
**➲ NATURE.COM**
For a slideshow of amazing Gemini images, see:
go.nature.com/h3kfvm

The turnaround is good news for astronomers of Gemini's seven member nations: the United States, the United Kingdom, Canada, Chile, Australia, Brazil and Argentina. Many are already planning to exploit the new Gemini multi-conjugate adaptive optics system (GeMS), an instrument that saw first light at Gemini South in December.

Adaptive optics are used at large ground-based telescopes to correct for image distortions caused by the motion of Earth's atmosphere. They monitor a bright guide star or backscattering from a laser beam fired by the observatory to measure the disturbances, then rapidly alter the shape of a mirror to correct them. But they sharpen only a small portion of the telescope's field of view. Astronomers have known for more than a decade how to make corrections over a larger area — by performing three-dimensional tomography of the sky with multiple lasers (R. Ragazzoni *et al. Nature* **403,** 54–56; 2000). But achieving this feat has been another matter.

## PAPER CHASE

On a per-telescope basis, the scientific productivity of the Gemini Observatory has lagged behind that of other large observatories of equivalent size.



SOURCE: D. CRABTREE

With its 50-watt laser split into five beams, Gemini has managed it first. The area over which it has achieved sharpness is more than 10 times that of other systems. "This is a big step," says Olivier Guyon, an adaptive optics expert at Japan's 8.2-metre Subaru Telescope in Hawaii. "It's going to be followed by everyone." Guyon adds that the system will help astronomers to study the skies more efficiently, by bringing large objects, such as nearby galaxies, into focus all at once, and by allowing more distant objects to be studied in a single snapshot.

Gemini South also unveiled a long-awaited spectrograph in December. This should allow users to calculate distances to some of the Universe's most remote objects, such as quasars, the luminous cores of distant galaxies containing giant black holes. Yet another instrument, an imager that could spot planets orbiting distant stars, is due by the end of the year. Gemini's notoriously clunky software interface for scheduling observations has also been overhauled.

Finding ways to connect with — or at least not alienate — Gemini's far-flung users is important, says Chaffee, because astronomers rarely fly out to the telescopes to conduct their observations. He is establishing a users committee that he hopes will represent observers' concerns, and a science and technology committee, which met last November, that will aim to help Gemini avoid repeating the costly and impractical instrument choices made in the past.

But Chaffee believes that instruments such as GeMS are already starting to remedy years of poor scientific performance. Virginia Trimble, an astronomer at the University of California, Irvine, says that, on a per-telescope basis, both publication and citation counts for Gemini have fallen behind those of its rivals (see 'Paper chase'), and the demand for telescope time has fallen. However, Gemini deputy director Nancy Levenson says that some of this decline in demand reflects waning interest by astronomers in the United Kingdom, which will end its membership of Gemini at the end of 2013.

Gemini's international agreement will be revisited in 2015. By then, Chaffee will have long since returned to retirement. His one-year contract is up in May, and the Association of Universities for Research in Astronomy, which manages Gemini, is due to name a new director within weeks. "I'm trying to leave the place in as good a shape as I possibly can," says Chaffee. ■

2&3 ILLUSTRATION

# Like it? Pay for it

## With conventional sources of money drying up, some scientists are turning to crowd-funding.

BY JIM GILES

In October 2010, Cesar Harada found himself in New Orleans with little money and a big idea. Harada, an engineer, had been working on oil-spill mitigation at the Massachusetts Institute of Technology in Cambridge. But he quit the lab in frustration at what he saw as a slow pace of work and a focus on expensive solutions. He travelled south to join the clean-up operation for the Deepwater Horizon oil spill in the Gulf of Mexico. Once there, his mind turned to a futuristic solution: a low-cost clean-up robot that local people could build and deploy themselves. Yet his two criteria for the project — a quick build and open-source intellectual property — all but ruled out academic or industrial funding.

Harada turned to Kickstarter, a website used by authors, film-makers and artists in search of project funding. He uploaded a pitch, set a goal of raising US$27,500 and listed a series of small rewards for donors. Then he started to network furiously. Money came in from friends and engineering colleagues. A few companies heard about his idea; they pitched in several thousand dollars each. Word reached people he had never met, and they contributed too. When Harada's funding appeal closed in April 2011, he had raised almost $34,000 — enough to assemble a team of engineers and build a prototype of the clean-up robot.

### PUBLIC INTEREST

If Harada's experience sounds like a one-off, think again. Crowd-funding — raising money for research directly from the public — looks set to become increasingly common. Established platforms such as Kickstarter are wooing scientists. And similar websites dedicated to connecting scientists with potential funders are being built, or have already launched. The public seems to be responding. Last year, for example, a group of scientists wanting to map water quality along the Mississippi River raised $64,000 in a trial project on an online crowd-funding platform called the Open Source Science Project (OSSP).

At a time when universities and research funding agencies are facing budget cuts, the strategy is attracting attention — as are other ways to raise philanthropic support (see page 254). "It's timely because of what's happening with traditional funding sources," says Daniel Gutierrez, co-founder of FundaGeek, a crowd-funding platform for technology projects that launched last month and is based in Yucca Valley, California.

For crowd-funding to make a real difference, advocates will have to prove that the process — which sometimes sidesteps conventional peer review — channels money to good projects, not just marketable ones. But if they succeed, there may be an unexpected bonus: it might help to forge a direct connection between researchers and lay people, boosting public engagement with science. "This is one of the most appealing aspects of crowd-funding," says Jennifer Calkins, an ecologist at Evergreen State College in Olympia, Washington, who has raised money for fieldwork on Kickstarter. "We can involve society in the creative journey that we make as scientists."

Online crowd-funding has already proved its worth in other fields. Kiva, a website through which individuals loan small amounts to entrepreneurs in the developing world, is one notable success: more than 600,000 lenders have channelled almost $275 million through the site since 2005. US President Barack Obama's 2008 election campaign raised a record-breaking $780 million, much of it from small online donations. And donors have pledged more than $100 million to 13,000 Kickstarter projects. By drastically simplifying the process of connecting donor with cause, the Internet has unleashed a new enthusiasm for giving.

Scientists have come a little late to the crowd-funding party, because they have conventionally had other funding streams. Jai Ranganathan, an ecologist at the University of California, Santa Barbara (UCSB), is one of several researchers trying to make up for lost time. Last November, he helped to launch the #SciFund Challenge, an exercise in which close to 50 research groups had six weeks to raise money through proposals on a crowd-funding platform called RocketHub, which mostly serves artists and entrepreneurs. The challenge raised a total of $76,000. Brian Meece, RocketHub's chief executive, based in New York, says that research projects are a "new and exciting" use for his platform, and that he will retain the science section now that the challenge is over.

## CASH FOR QUESTIONS

Other crowd-funding enthusiasts are developing donor sites dedicated exclusively to research projects. Sixteen projects are currently vying for funds on SciFlies, a site launched last November by David Fries, a marine engineer at the University of South Florida in St Petersburg. This year, the OSSP hopes to follow up on its success with the Mississippi study by launching a fund-raising appeal for around ten research proposals, says Priyan Weerappuli, a neuroscientist at the University of Michigan in Ann Arbor and the founder of the project.

Each site operates in a slightly different way, but there are common themes. Researchers start by describing and pricing a project, which they submit to the site for approval. If accepted, the pitch is placed online and donors have a few weeks or months to read the proposal and make a donation. Some sites operate on a non-profit basis and channel all proceeds to researchers; others are commercial concerns and take a cut of the money raised.

But although cash-starved scientists are lining up to list their projects, some are also expressing concerns. Take the issue of peer review. SciFlies and the OSSP post projects only after passing them through an expert review process, but Kickstarter's only requirement is that projects have "a creative purpose" — as defined by the site's owners. Projects in the #SciFund Challenge did not undergo formal peer review: Ranganathan and co-founder Jarrett Byrnes, a fellow UCSB ecologist, checked only for obvious fraud. "I don't care if people have badly thought-out projects," says Ranganathan.

**➲ NATURE.COM**
**Tell us your crowd-funding tips by commenting at:**
**go.nature.com/a3bi7o**

That may sound like a recipe for shoddy science, but crowd-funding advocates say that the process has an inbuilt peer-review system, driven by the donors. Most donors will hear of a project through their social networks. They might be former colleagues of the project owner, or members of the public interested in an ecological study site. So project owners put their reputation among their peers and supporters on the line every time they post a proposal.

"There's a strong incentive to be honest," says Kickstarter co-founder Yancey Strickler. "Social forces carry a lot of weight." The system also puts a premium on inventive, well-thought-out proposals. A poorly conceived pitch that attracts no funds will do nothing for a scientist's career; nor will one that never delivers on its promises. "It may not be formal peer review, but crowd-funding has validation based on common trust," says Meece. "It's a pretty heavy filter." Even Sally Rockey, deputy director for extramural research at the US National Institutes of Health in Bethesda, Maryland, sees benefits in an alternative evaluation system if it helps organizations to achieve their research goals. Peer review "is not the only model", she says.

## MASS APPEAL

### *How to woo the crowd*

The owners of crowd-funding sites give their tips on pitching winning scientific proposals.

- Create a compelling story about your research. Who will it benefit? And how? Then tell that story to camera — many sites allow project owners to upload short videos as part of the pitch.
- Devise clever rewards for donors. Think about giving away T-shirts decorated with project logos or, for big donors, a chance to visit your lab. Most sites require project owners to offer some reward, but bear in mind the time and expense required to produce and distribute whatever you offer.
- Use your social network, online and offline. Tell friends about the project, and ask them to tell their friends. Tweet it, blog it, publicize it on Facebook.
- Study previous successful pitches. Talk to the researchers behind them. Learn what works and incorporate it into your pitch. **J.G.**

Some sites are trying to enhance this informal review process. FundaGeek has a discussion forum, the 'Geek Lounge', where potential donors are encouraged to debate the merits of a proposal. Last August, the equivalent forum on Kickstarter helped to halt one questionable project. The proposal, for a product called the Tech-Sync Power System, aimed to develop a smartphone app that controls home lighting. It attracted more than $27,000 in pledges, but Kickstarter users with electronics knowledge started to question the viability of the system. The project owner, who could not be reached for comment, eventually deleted his proposal as the criticism mounted, and none of the donors lost their money.

## THE HARD SELL

Another objection to crowd-funding may be harder to shake. To sell a project, researchers need an attention-grabbing story (see 'How to woo the crowd'). That is easy to construct if your subject of study is, say, saving pandas or curing cancer. It is less so for researchers working on polymers. So will crowd-funding prove profitable only for 'sexy' science?

Fries concedes that crowd-funding inherently favours certain types of project, particularly those in applied research. He is an optimist, saying that if the approach takes off, conventional funding agencies will simply have to compensate by upping their support for basic science.

Ranganathan, an enthusiastic communicator who runs his own podcast, bristles at the suggestion that crowd-funding will create a two-tier system. "It's all about telling a compelling story about the research," he says. "Panda researchers start ahead, but I 100% believe anyone can do it." A polymer chemist might, for example, focus on new materials that could come out of his or her work.

The pressure to communicate the potential fruits of a research project should not be seen as a burden, adds Ranganathan. Most crowd-funding sites expect project leaders to offer donors something in exchange for their contribution, such as regular updates on the progress of the research. For those who make larger donations there might be visits to a lab or field site. In the case of the Mississippi water-quality study, donors in the region were encouraged to help with collecting water samples from the river. This process should help to forge stronger bonds between researchers and the public.

Whether all this works in wider practice remains to be seen, but many welcome the experiment. "Science thrives on diversity," says Jack Stilgoe, who studies science and society at the University of Exeter, UK. "We shouldn't be afraid of innovations in how it is funded. We should be more afraid when research money is all getting spent in the same way on the same sorts of things." ■ **SEE EDITORIAL P.238 AND COMMENT P.260**

**Jim Giles** *is a freelance writer based in San Francisco, California.*

YOUR LAB
HERE

# Sponsor my science

Philanthropists will sometimes give large sums of money to support science — but researchers have to learn how to sell themselves first.

**BY HEIDI LEDFORD**

2&3 ILLUSTRATION

Asking someone to give you a million dollars is not easy. In 1995, Bruce Walker was seeking philanthropic support to expand his HIV research programme at Massachusetts General Hospital in Boston. He had identified a possible donor: a venture capitalist and the brother of one of his patients. He had teamed up with a personal coach in fund-raising and rehearsed his pitch over and over. But during lunch with the prospective donor, he still nearly choked. "I couldn't quite get it out to say, 'Would you give us a million dollars?'," Walker recounts.

Eventually he spat out his request — and was astonished when the venture capitalist agreed to the entire sum. "At that point, I fell off my chair," Walker says. He then raised another US$100 million from philanthropists Terry and Susan Ragon to launch the Ragon Institute of MGH, MIT and Harvard in Charlestown, Massachusetts, in 2009. He and his colleagues are now planning a formal seminar series on raising philanthropic support. His success at fund-raising, he says, "is not because there's something special about me. It's because I got a little bit of training and I put in a lot of effort."

Walker is not the only one with a determined approach to philanthropy. As public funding for research dwindles, scientists are increasingly seeking private benefactors (see page 252). But they have a lot to learn if they are to win trust and money: how to schmooze contacts, promote their science and deliver results to deadline — all without over-promising on the work (see 'How to woo philanthropists'). Much of this does not come naturally to scientists. "You have to sell yourself," says Cheryl McEwen, who, with her husband Rob, donated US$10 million to found the McEwen Centre for Regenerative Medicine in Toronto, in 2003. "But if you can build a case that we can understand, we'll be there for you."

➲ **NATURE.COM**
Read a Comment on the dangers of philanthropy:
go.nature.com/te4dz8

To help build that case, research institutions — particularly those in the United States — are becoming savvier about how to approach philanthropists. In April 2010, Steve Rum, the vice-president for development and alumni relations at Johns Hopkins School of Medicine in Baltimore, Maryland, started training faculty at the school in fund-raising techniques. Elsewhere, scientists are enrolling in classes provided by external institutes. Researchers are happy to have the help, and not just because it can bring in new income. It can also allow them to pursue projects that government funding committees would find too risky to support. "When you come to the government agencies with their committees and extreme accountability and taxpayer dollars, everything tends to come out a bit plain vanilla," says Tim Hunt, a Nobel-prizewinning biochemist at Cancer Research UK in London, who credits philanthropic support with helping him to continue his basic cell-biology studies even when applied research was more in vogue. "It's valuable to have funding that's a little bit out of the mainstream."

## PERSONAL APPROACH

Philanthropy is supporting a growing slice of science: donations from US foundations to science, technology and medical research have grown from $793 million in 1999 to $1.7 billion in 2010, according to figures from the Foundation Center, an organization based in New York that analyses information about philanthropy. But although many scientists have experience in sending grant applications to foundations, few try to win multimillion-dollar donations from rich and generous individuals.

Convincing a philanthropist to fund a programme means cultivating a personal relationship with them. "It's an old adage, but it's completely true: people give money to people," says Thomas Pierson, chief executive of the SETI Institute, an astrobiology research centre in Mountain View, California. After SETI was cut loose from NASA in 1993, it relied on Bernard Oliver, a well-connected member of its board and a former head of research and development at the computer firm Hewlett Packard in Palo Alto, California, to take its case to local entrepreneurs. The result: support from a who's who of technology luminaries, including Microsoft co-founder Paul Allen.

Most scientists will need to look outside their usual network of colleagues and supporters for prospective donors, says Robert Klein, a real-estate magnate who helped to raise $34.5 million in the 2004 campaign for California's $3-billion stem-cell initiative. Most of that, he says, came from wealthy families with a vested interest — usually a sick relative — in seeing stem-cell technologies succeed. To reach these families, Klein recommends first building contacts with patients and those who support them. "Have a patient advocate call ahead, and sit next to you while you make the pitch," says Klein. "Better yet, if you can get them to do it, have the advocate make the pitch for you."

Walker urges researchers to be on continual alert for new contacts. "When you travel, always try to get upgraded to business class," he says. "Then talk to the people around you about what you do."

This type of schmoozing takes time and a certain personality. Rum says that Morton Goldberg, an ophthalmologist at Johns Hopkins, is a fund-raising "superstar" — helping to build an endowment that now funds more than 30 professorships. The secret to his success, says Rum, is his willingness to forge lasting relationships with philanthropists. Goldberg says he considers many of his donors to be close friends. He has travelled with them, spending weekends on their boats and in their vacation homes. "It has been an extremely rewarding experience, far beyond any connection with money," he says.

Many physicians are reluctant to solicit their former patients, and with good reason, says bioethicist Sheldon Krimsky of Tufts University in Boston. "It's not OK for a doctor who is actively treating a patient to do it." Even after a patient has left the hospital, he or she may return, says Krimsky, and a physician who has received that patient's support may now be expected to provide special treatment.

Another pitfall of philanthropy, Krimsky says, is donors who attach strings to their gifts. When the Charles G. Koch Charitable Foundation in Arlington, Virginia, agreed to fund a faculty position in economics at Florida State University in Tallahassee, it demanded the right to determine the criteria used to pick a professor, and to veto candidates it did not like (S. Krimsky *Nature* **474,** 129; 2011). "It was so egregious," says Krimsky. The university accepted the terms in 2008, but has argued that input from the foundation during the hiring process has not compromised its academic integrity.

> *"Try to get upgraded to business class."*

Krimsky worries that tough economic times can make researchers and their institutions more willing to accept intrusions into academic freedom in exchange for funding. Scientists agree, but several who have received philanthropic support point out that they appreciated the sound financial and strategic advice the donors offered.

For many scientists, the most difficult step is pitching their work as if they were an entrepreneur in need of investment, says Garen Staglin, who, with his wife, Shari, donated some of the family's earnings from their vineyard in Napa Valley, California, to mental-health research after their son was diagnosed with schizophrenia. "You have to go into sell mode," he says. "You have to be able to say 'there are no guarantees here but if this works, this is the kind of stuff we think is possible in our lifetime.'" But that type of sell tends to be easier for applied projects, such as potentially life-saving medical studies, than for basic research.

Selling research can also require a willingness to set deadlines. "In difficult economic times, the philanthropies become very goal oriented," Klein says. It's a practice that many scientists resist, because science rarely goes according to plan. When it doesn't, fund-raising specialists say that it's best to acknowledge the failure and stress the importance of the lessons learned. "You will still give them the confidence that there are incremental improvements," says Klein.

## FUND-RAISING 101

### *How to woo philanthropists*

Advice from successful fund-raisers on how to sell science.

- Network outside your usual circles: look for contacts — and contacts of contacts — who can strike an emotional chord with philanthropists.
- Learn how to frame your research and its importance in a few sentences. Practise this 'elevator pitch' in front of the mirror, and with anyone who will listen.
- Aim high: look for hints about how much a donor might be willing to commit.
- Establish deadlines: break projects into small chunks and assign completion dates for each step. Explain the vagaries of scientific research to your donor and don't be afraid to adjust timelines.
- Reward your investor: send regular progress updates, acknowledge their sponsorship in publications and presentations and encourage contributors to come to the lab for visits and tours. **H.L.**

### PHILANTHROPY SCHOOL

Research institutions often rely on professional fund-raisers to approach philanthropists, but Johns Hopkins isn't alone in training faculty members. Advancement Resources in Cedar Rapids, Iowa, has seen a fivefold increase in demand for its fund-raising training over the past four years, says chief executive Joe Golding. Institutions pay $17,500 for a four-hour workshop of lectures and role-playing exercises.

Golding says that basic researchers would do well to target entrepreneurs. "They are, by definition, risk takers," he says. "And they value leverage." Stress that a $30,000 investment now could lead to a million-dollar government grant later, he advises, and emphasize the impact a project would have on a topic that is close to the entrepreneur's heart. Golding, Rum and other experts also tell their trainees to listen for clues to how much a donor might be willing to commit. "If I ask for $100,000 and they say 'yes' right away, then I didn't ask for enough," says one fund-raiser at a large research university who asked not to be named. "It's a common mistake."

Rum even conducted a trial to find the best way to teach faculty members about how to approach former patients. He divided faculty members into three groups. One received only weekly fund-raising lessons by e-mail. The second also attended a lecture on the topic. The third received individualized coaching, including how to prepare an 'elevator pitch' — a brief description of research so compelling that it could capture someone's attention in the time it takes to travel between floors. Three months after coaching, the first two groups had brought in no money. The third had collected five gifts totalling $219,550 (S. Rum and S. M. Wright *Acad. Med.* **87,** 55–59; 2012).

Walker says that this type of personal coaching — not to mention the agony of asking for money — has been more than worthwhile. "I think we as a scientific community don't do a great job in articulating the transformational power of philanthropy," he says. "But that funding has allowed me to do things that I never could have considered if I'd been trying to do it through traditional sources."

Too few researchers take advantage of philanthropy, says Goldberg. "The most common mistake is not to try." ■ SEE EDITORIAL P.238 AND COMMENT P.260

**Heidi Ledford** *is a reporter for* Nature *based in Cambridge, Massachusetts.*

# nature|methods

## Genome Editing with Engineered Nucleases

January 2012 | volume 9 | number 1

nature|methods

www.nature.com/naturemethods

Techniques for life scientists and chemists

- A rhodopsin-based voltage indicator
- A fluorescent molecular tension sensor
- High-resolution 5-hydroxymethylcytosine detection
- A Review of genetic tools for Drosophilists
- METHOD OF THE YEAR 2011

## Method of the Year

Nucleases – DNA cutting enzymes – can be designed to cut genomic DNA at a desired sequence, so that researchers are increasingly able to precisely modify specific genes in several species. The two main classes of enzymes that can be used for this purpose are the zinc finger nucleases (ZFNs), which have been developed over the past two decades, and the newly reported transcription activator-like effector nucleases (TALENs). These tools can be used to tinker with genes in many ways – to delete, mutate, modify or tag them in their endogenous genomic context. Both laboratory design methods for these enzymes as well as commercial versions are becoming steadily more accessible to the wider research community; genome editing with engineered nucleases is thus poised to change the way scientists study gene function.

*Nature Methods* presents a series of articles that report how these technologies developed, the difficulties in designing good tools, and their exciting applications. The Methods to Watch feature provides a look at possible future Methods of the Year.

Access the **Method of the Year** and **Methods to Watch** by visiting:
**www.nature.com/nmeth/focus/moy2011**

nature publishing group npg

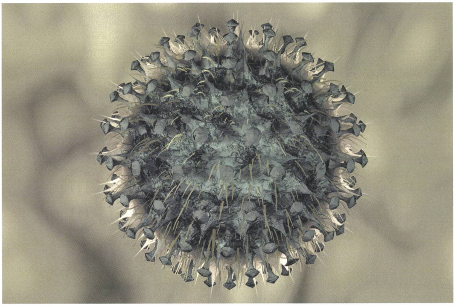
# COMMENT

**PHILANTHROPY** Charities should pay their fair share of research costs **p.260**

**VACCINES** More collaboration would spur biomedical research innovation **p.261**

**EXHIBITION** The scientific legacy of Robert Scott's ill-fated expedition **p.264**

**OBITUARY** Paul Doty, chemist and arms-control advocate **p.266**



OXFORD SCIENTIFIC/PHOTOLIBRARY/GETTY

# The fight over flu

A proposal to restrict the planned publication of research on a potentially deadly avian influenza virus is causing a furore. Ten experts suggest ways to proceed.

## RON FOUCHIER & AB OSTERHAUS

## Globalize the discussion

*Erasmus MC, Rotterdam, the Netherlands*

So far, most of the human deaths from the deadly H5N1 strain of bird flu have occurred in Asia and the Middle East. Many labs worldwide — including ours — are trying to understand what makes the virus so virulent, and how to stop it. H5N1 research is thus a global issue, yet the entire research community seems to be following the advice of one country.

We are not questioning the unprecedented recommendations last month from the US National Science Advisory Board for Biosecurity (NSABB) to remove key details from the methods and results sections of published papers, including our own, submitted to *Science* (see *Nature* **481,** 9–10; 2012). But we do question whether it is appropriate to have one country dominate a discussion that has an impact on scientists and public-health officials worldwide. This discussion should include the perspective of people in regions where H5N1 has infected humans. Will the NSABB also advise on which international researchers and officials have the right to see the full papers, to help implement urgently needed surveillance and other intervention strategies?

It is not clear whether an international discussion would lead to different recommendations. There is no global equivalent of the NSABB, but many European experts that we have seen quoted in the press believe that the research should be published in full. We don't know the worldwide opinion until a group of experts from all parts of the globe is formed. An issue this big should not be decided by one country, but by all of us.

## JOHN STEINBRUNER

## A system for redacted papers

*Director, Center for International and Security Studies at Maryland, University of Maryland, College Park*

If the two papers (submitted to *Nature* and *Science*) describing a transmissible form of the H5N1 virus are the first to be published with key details missing, they probably won't be the last. We need to establish both a short-term and a long-term solution for how the scientific community should handle such a publication. Who decides who should have access to the full details? Who monitors the community so that the details don't get passed around outside the group of experts cleared to receive them?

I believe that the entire process must be regulated by a global health body, ideally the World Health Organization (WHO). Already, a WHO committee oversees all research involving the smallpox virus. A similar, more developed system could work for H5N1 and other deadly pathogens. An international group of experts would approve research involving those agents, decide who will have access to the details of papers that come out of that work, and hold those vetted individuals accountable for what they do with the information. For instance, such a system might allow permitted experts to view papers only electronically, so nothing is on paper. A database could record the privileged few people who have seen the full paper, and what they do with the information. In all likelihood, legal safeguards would be needed to protect the rights of people who receive the sensitive information. Such a process would not offer complete protection against misuse, but it would show the scientific community that the committee is watching what they do.

Because dangerous pathogens are a global issue, any procedure would need buy-in ▶

▶ from all countries, who would have to give the committee binding jurisdiction over research involving extremely dangerous agents. This will take some time. In the short term, the WHO or some other global health organization should immediately establish an ad hoc committee to review who should receive access to the full H5N1 papers, and ensure that the details do not circulate widely.

Most importantly, these discussions should not be controlled by officials focused on national security. H5N1 is primarily a matter of public health. If there is a threat of bioterrorism, let it be judged by a global health organization, which can set rules that do not deprive scientists of information that could save millions of people in the case of a natural pandemic. If national-security organizations become involved, they will vet scientists on the basis of citizenship, and will be inclined to discriminate against those countries in which terrorists have found refuge. But some of those countries are among the few that have experienced human deaths from H5N1, and are most likely to witness the origins of a natural pandemic. It is crucial that scientists and other experts are judged on their qualifications, not on their nationality. If the world is to accept the process, national-security officials cannot be allowed to dominate the discussion.

### KWOK-YUNG YUEN

## The Hong Kong perspective

*Chair of Infectious Disease, Department of Microbiology, University of Hong Kong*

As a scientist working in Hong Kong — the site of the first human epidemic of infection by the highly fatal H5N1 virus — I appreciate the public-health significance of knowing which mutations confer airborne transmissibility in an animal model. The new, much-debated research provides this information. Finding similar genomic signatures in animal or human viruses collected from the WHO Global Influenza Surveillance Network may alert public-health workers to an impending epidemic of unthinkable magnitude or severity. But I also appreciate the possibility that such mutants could cause a global disaster if accidentally or deliberately produced and released into animal and human populations. Consequently, I support the recommendation from the NSABB to remove key details from the papers describing this work.

Biological warfare is familiar to people living in this part of the world. During the Sino-Japanese War in the 1930s and 1940s, scientists and physicians of the Japanese army's Unit 731 infected civilians and prisoners of war in Manchuria, China, with *Yersinia pestis*, the bacterium behind the Black Death, which was first identified in plague patients in Hong Kong in 1894. When the unit found that the bacteria caused severe organ damage when serially inoculated into prisoners of war, the unit began spreading plague-contaminated fleas in China, causing outbreaks.

N. SHRESTHA/EPA/CORBIS

**South Asian countries such as Nepal are among the most likely to be hit by avian flu epidemics.**

Censoring scientific data for publication will not stop rogue individuals or nations from developing a deadly and highly transmissible form of H5N1, but it would at least buy some time to find and stockpile the appropriate antivirals, immunomodulators and vaccines to protect against most variants of H5N1. Even if the publications omit the methods for making such a deadly virus, the genomic signatures associated with airborne transmissibility should be known to the directors of all public-health laboratories in the WHO surveillance network, after they sign an agreement of confidentiality.

### D. A. HENDERSON

## The ultimate biological threat

*Center for Biosecurity, University of Pittsburgh Medical Center, Pennsylvania*

The H5N1 influenza strain poses a potential biological hazard far more serious than any we have ever known. It is a virus that is capable of killing half its victims, a proportion greater than that for any other epidemic disease. Were that coupled with the transmissibility of a pandemic flu virus, it would have characteristics of an ultimate biological weapon unknown even in science fiction (see *Nature* **480,** 421–422; 2011). We should not publish a blueprint for constructing such an organism.

### LYNN KLOTZ & ED SYLVESTER

## Worry about lab infections

*Center for Arms Control and Non-Proliferation, Washington DC; Walter Cronkite School of Journalism at Arizona State University*

Asian bird flu is just one of the extremely dangerous pathogens researched in laboratories throughout the world. Along with the two labs that created a potentially contagious form of H5N1, at least 40 others worldwide investigate deadly, highly contagious pathogens not currently present in human populations, such as the SARS virus and the recently resurrected 1918 pandemic flu virus. Public-health experts worry about natural pandemics, and governments worry about the risk that these pathogens pose to national security — but the probability of accidental release is likely to be much higher.

We have analysed the likelihood of escape from 42 labs, using 1% as the estimated probability of an escape from a single lab in a single year. This approximates the historical probability, obtained by dividing the documented number of escapes of these pathogens (3, each involving the SARS virus) by our estimate of the total number of lab-years of research on these pathogens since 2003 (more than 300 lab years). Lab infections can easily spread: in 2004, after the only natural SARS outbreak in humans was contained, two graduate students lab-infected with SARS in Beijing infected seven others, causing one death.

Our analysis shows that the probability of an escape from at least one of 42 labs in a single year is 34%; within less than 4 years, the chance of escape reaches 80%. If more labs begin studying the (potentially) human contagious form of H5N1, that interval will decrease even further.

In comparison, the average time between the three natural influenza pandemics in the last century was about 30 years (1918, 1957 and 1968). We are creating a risk that is much greater than that posed by nature. Laboratories need stronger precautions, such as mandatory quarantine of lab staff working on live dangerous pathogens until they can be certified clear of infection. Regulators should not be sitting idly by, while the threat of a man-made pandemic looms.

## JEFFERY K. TAUBENBERGER

## Study how viruses swap hosts

*Chief, Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, US National Institute of Allergy and Infectious Diseases*

How influenza viruses adapt to humans and cause disease has been a significant interest of my laboratory, including sequencing the genome of the 1918 H1N1 influenza, the virus that caused the worst pandemic on record. This work has led to research on how influenza viruses switch hosts and cause severe disease, and is serving as a basis for new vaccines and therapies. As a result of these studies, scientists and public-health officials are better prepared to prevent or mitigate a similar future pandemic. Studies examining adaptation and transmission of avian H5N1 influenza virus are similarly important.

For reasons not yet clear, some influenza viruses adapted to poultry can acquire sets of mutations that kill almost every bird that they infect, but rarely infect humans. Since 1997, H5N1 viruses have devastated poultry flocks in southeast Asia and other parts of the world, and also caused disease in a limited number of the presumably tens of thousands of humans who have had direct contact with infected birds. Although human infections are uncommon, the high fatality rate in those people infected by H5N1 — more than 50% — raises serious concerns.

A growing body of evidence suggests that influenza host-switching processes are complex, and may be unique to each virus. Changes in the 1918 H1N1 virus crucial for adapting to humans, for example, were not found in the 2009 pandemic H1N1 virus. It is essential to continue research into this with H5N1 and other emerging pathogens, and to investigate how they adapt and cause disease. The work could provide better surveillance tools to detect viruses that are in the process of adapting to humans. Scientists must be able to continue to work with these viruses safely and appropriately, in multiple teams utilizing the talents of many researchers, and under the numerous layers of protection and oversight that safeguard scientists and the public.

➲ NATURE.COM
See *Nature*'s web focus on the 1918 influenza pandemic:
go.nature.com/9rcpjv

Of course, scientists will have to continuously re-examine and discuss the balance between the benefits of conducting and publishing pathogen research, and the risk of misuse of such knowledge. We should continue the dialogue as we continue the research.

## RICHARD H. EBRIGHT

## Mitigate the risks of release

*Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey*

Engineered derivatives of highly pathogenic avian influenza virus capable of aerosol transmission in non-rodent mammals pose potential threats to human health and to food production. The major concern is accidental release, for example through infection of a lab worker who then infects others. Deliberate release by a disturbed or disgruntled lab worker, bioterrorism and biowarfare are also concerns.

The following steps should be taken immediately. To address accidental release, assign the viruses as pathogens that require the highest biosafety level ('level 4', not 'level 3+' as has been the case to date). To address deliberate release, assign the viruses as pathogens that require the highest level of biosecurity ('Tier 1', like smallpox virus and anthrax bacterium, under the revised US select-agent rule that enters into force this year).

To minimize risks from future research, additional steps should be taken. We need to implement a system of mandatory prior review of research directed at increasing a pathogen's virulence, transmissibility or ability to evade countermeasures. We also need to re-evaluate, and preferably terminate, biodefence expenditures on research directed at creating and assessing new biothreats, as opposed to addressing existing biothreats. Creating and assessing new threats rarely increases security. Doing so in biology — where the number of potential threats is nearly infinite, and where the asymmetry between the ease of creating threats and the difficulty of addressing threats is nearly absolute — is especially counterproductive.

## DAVID L. HEYMANN

## We will always need vaccines

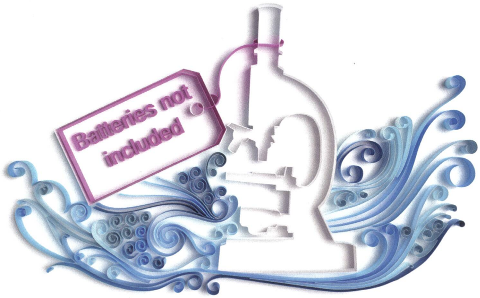
*Head, Centre on Global Health Security, Chatham House, London*

The creation *in vitro* of highly transmissible forms of H5N1 is a fresh reminder of how the endgame in infectious-disease control has changed.

For years, scientists and officials have believed that the ultimate application of disease control was eradication. In a world free from a particular pathogen, the disease no longer occurs, and the expense — and small risk — of vaccination is avoided. As a result, there would be no need to develop and maintain vaccine stockpiles. This idea began to change in 2002, when the poliovirus was synthesized *de novo*. The change has been profound — the endgame must now take into consideration the possibility that consolidation and safe storage, or destruction, of remaining polioviruses is not sufficient to eliminate the risk that it will be reintroduced. With the possibility of synthesizing a virus, it is likely that we will always need to have vaccines readily available.

But some say that maintaining samples of deadly pathogens is a necessary precaution. Take the discussions about the samples of smallpox virus (*Variola*), maintained at two WHO-sanctioned laboratories since the virus's eradication. Some now argue that we need these samples to study the virus, and to develop new vaccines and therapies in case of a bioterrorist attack. Indeed, the WHO Advisory Group of Independent Experts that reviews the smallpox research programme noted this year that DNA sequencing, cloning and gene synthesis could now allow *de novo* synthesis of the entire *Variola* virus genome and creation of a live virus, using publicly available sequence information, at a cost of about US$200,000 or less.

It is tempting to argue that, for the sake of public health, research that constructs highly pathogenic forms of viruses, such as H5N1, should be squashed. But even if we eliminated H5N1 and destroyed all lab samples, we would still need to consider keeping and producing a vaccine, because the virus could potentially be recreated. And stopping this type of research would open a dangerous door. Although adding and deleting genes can create super-strains that put the entire world at risk, such research also helps to develop public-health tools such as vaccines and diagnostic tests. Preventing this research would also prevent us from using all possible scientific options to prepare for naturally occurring — or deliberately caused — outbreaks. ■



2&3 ILLUSTRATION

# The price of charity

Philanthropists should pay their fair share of research costs, says **Patrick Aebischer**.

With public funding for universities diminishing, private and philanthropic sources are increasingly being pursued to support academic research. Although welcome, this money comes with a catch: charities and foundations rarely pay the full costs of running a lab — building construction, maintenance, utility bills or salaries, for instance — and tend to be selective in the projects they sponsor. Universities therefore need to do more to pass on the true costs of research to donors. They must also have a balance of income sources to avoid bias in their research directions.

Private foundations have been significant sponsors of science since the nineteenth century. Until the US National Institutes of Health started funding extramural research in the mid-1940s, for example, more than one-quarter of US medical research was paid for by philanthropic organizations. As public funding grew, the share contributed by foundations and charities declined to less than 4% by 2007 (ref. 1).

Since then, support from non-profit organizations — especially those associated with rich individuals such as Bill Gates — has risen worldwide. In Europe, philanthropic sources now supply 6.5% of competitive research funding on average — 3–4% in most European countries and almost 10% in the United Kingdom[2]. At my own institution, the Swiss Federal Institute of Technology in Lausanne, private sponsorship has tripled, from 3% to 9% of research income over the past 10 years.

Biomedicine is benefiting most. And nowhere more so than in the United Kingdom: led by the Wellcome Trust, charities accounted for 40% of British biomedical research funding in 2009 (ref. 3).

## LEGACY OF URGENCY

Private institutions, however, have a disproportionately large influence on the research landscape relative to their financial contribution. Because they can make funding decisions quickly, their money is welcomed by principal investigators looking to fill funding gaps. But directed awards can divert an entire group's research towards one end.

Keen to ensure that results are delivered and milestones met, charities' legacy of urgency has taken them away from backing long-term research projects and towards satisfying narrow goals that fit their mission. Since the boom in 'venture philanthropy' in the 1990s, the personal agendas of powerful entrepreneurs have accelerated research into AIDS; infectious diseases such as malaria and tuberculosis; cancer; and neurodegenerative diseases such as Parkinson disease.

Charities often send the message that not one cent invested is spent on anything other than finding the cures and so minimize their contributions towards a university's overhead costs. That means that institutions with many privately funded projects are effectively 'punished' for their success. To meet the higher research-infrastructure costs, universities may drain resources from education, or diminish 'expensive' disciplines such as physics, chemistry or engineering, in which philanthropic support is scarce.

Private bodies should not hijack university resources. They should contribute a fair share of the expense of a sustained research enterprise. To make it easier for them to do so, universities should better identify the full cost of research activities and pass it on. And because most charities operate internationally, these overheads should be harmonized worldwide.

## FULL COSTS

Such an accounting effort is ongoing in some countries, especially the United States. But the concept of overheads is almost unknown in many European countries, where universities are largely state supported. Some European universities are addressing this issue, in response to requirements by the European granting system to supply full research costs.

Estimates of overhead costs vary by field and by country. In the United States, they typically range from 40% to 70% of grant income, depending on whether the principal investigator's salary is included. But awards are usually lower in Europe. For example, in 2008, the Swiss Parliament agreed to compensate universities with an overhead of approximately 10% of each successful project submitted to the Swiss National Science Foundation. Grants from the Seventh Framework Program and the European Research Council support a maximum overhead of 20%. This is inadequate and should rise.

Universities can do much to manage their varied funding streams. They should develop, champion and apply transparent full-cost accounting mechanisms. Leaders engaging in fund-raising should promote the inclusion of overhead costs. And academies must avoid propping up underfunded research activities with educational resources.

Ultimately, universities should not rely on short-term funding for basics such as infrastructure and faculty salaries. They need to retain a healthy level of public funding if they are to survive long term. Universities have a mission of education and of long-term, fundamental research; charities aim to solve real-world problems now. Fruitful partnership should imperil neither but achieve the noble goals of both. ■

---

**Patrick Aebischer** *is president of the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
*e-mail: patrick.aebischer@epfl.ch*

1. Dorsey, E. R. *et al. J. Am. Med. Assoc.* **303,** 137–143 (2010).
2. Estermann, T. & Pruvot, E. B. *Financially Sustainable Universities II: European Universities Diversifying Income Streams* (European University Association, 2011).
3. Martin, L. *All Together Now: Improving Cross-sector Collaboration in the UK Biomedical Industry* (NESTA, 2011).

D. LEFRANC/GAMMA-RAPHO/GETTY

DAILY NEWS
NBC NEWS STAFFER STRICKEN
ANTHRAX HERE
HOW TO PROTECT AGAINST THREAT
Brokaw's aide recovering; powder in letter suspected
New York Times evacuated after getting similar letter
FBI probing St. Petersburg, Fla., postal connection
Rudy to city: Stay calm
FULL COVERAGE BEGINS ON PAGES 2 & 3
I ❤ NY

**Despite threats such as the 2001 US anthrax attacks, vaccine innovation has slowed in the past 40 years.**

PUBLIC HEALTH

# Biodefence built on teamwork

The golden era of US vaccine research holds beneficial lessons for today, finds **John Grabenstein**.

Biomedical researchers who are familiar with the US National Institutes of Health as a funding body and agenda-setter might be surprised by *Long Shot*. Science historian Kendall Hoyt shows that there was once a different way to drive medical progress. In the 1940s and 1950s, focused expertise and goal-oriented collaboration between the military, academia and industry yielded key vaccine innovations that bolstered national security, yet had benefits for civilians. That era of medical discovery stands in marked contrast to today's patent constraints, regulatory hurdles and disjointed efforts.

The book is timely, with the United States and other countries continuing to search for pharmacological defences against unforeseen biological threats such as the 2001 US anthrax attacks, which killed 5 people and infected 17. The current US biodefence effort has been expensive — less so when compared with the cost of certain military aircraft — yet it has not focused adequately on developing products that are usable. Recalling a time when introducing vaccines seemed easier could provide salutary lessons for today.

➲ NATURE.COM
For more on civilian and military disease collaborations:
go.nature.com/thfn8s

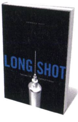
Hoyt begins her deft treatment by describing the threats from pathogens and how she measures vaccine innovation, then reviews vaccine research during and after the Second World War. She also discusses recent factors that have disrupted vaccine-producing government–industrial networks and led to contemporary frustrations.

In the 1940s, vital research aiming to create practical products for protecting soldiers and nations overwhelmed theoretical studies. Most developments were not a response to enemy bioweapons, but rather to viral and bacterial pathogens in countries where troops were based, or to diseases linked to crowding or poor hygiene. For example, influenza A and B vaccine was developed in response to the 1918–19 pandemic that killed tens of millions of people; pneumococcal polysaccharide vaccine came about because of the crowded conditions in barracks that often led to pneumonia; and typhus and tetanus vaccines were produced to combat two common diseases faced by troops throughout history.

As Hoyt shows, the urgent need to find vaccine solutions drove highly effective collaborations between various military, academic and industrial research bodies. But that urgency led to stumbles too. When production of yellow-fever vaccine was scaled up to enable vaccination of US Army troops in the 1940s, the vaccine was stabilized with human serum that was unknowingly contaminated with hepatitis B virus, leading to almost 50,000 hospital cases. Formulation changes later allowed the vaccine to be used to prevent yellow fever across broad swathes of the globe.

The book's narrative is enriched by a description of how the US Office of Scientific Research and Development, the Office of the US Army Surgeon General and the War Research Service conducted vaccine-research planning and operations during the war. With peace in 1945, wartime scientific relationships continued to bear fruit, such as vaccines against adenovirus and meningococcal infections. She lauds collaborations throughout the 1960s between the Walter Reed Army Institute of Research in Silver Spring, Maryland, and vaccine firm Merck, Sharp and Dohme in West Point, Pennsylvania. She describes them as science integrators that brought together experts from disciplines including medicine, immunology and virology, all focused on common goals. Much of their success, she says, is due to that focus and to teamwork that avoided over-specialization.

To measure vaccine innovation between 1900 and 1999, Hoyt reconstructs historical records of annual US vaccine introductions. This was a challenge: responsibility for vaccine registration (or licensing) was transferred from bureau to bureau. Between the establishment of the US Hygienic Laboratory in 1902 to that of the Center for Biologics Evaluation and Research in 1987, there were seven institutional transitions, and multiple variations in record-keeping.



**Long Shot: Vaccines for National Defense**
KENDALL HOYT
*Harvard University Press: 2012. 320 pp. $29.95, £22.95*

According to Hoyt, vaccine innovation activity peaked in the 1940s at 50 innovative licences per decade, remained at

35–40 licences per decade through the 1960s, and declined steadily to 12 per decade in the 1990s. The numbers reflect all vaccines developed for endemic and exotic pathogens. She convincingly contests similar analyses.

Today's biological revolution cuts both ways, offering great advances in medicine while providing new means of attack for terrorist groups. It is thus short-sighted that US biodefence funding is largely overseen by congressional committees that are oriented towards health rather than national security.

If the 1940–69 model was a worthy paradigm, why did we migrate away from it? Hoyt attributes the change to several factors. Government contracting and licensing became mired in bureaucracy, and concerns over intellectual property curtailed collaborations. Oversight of vaccine development across a number of bodies became inadequate. The political constituency for biodefence issues weakened and funding for products suffered. Arguably, the lack of US government prioritization of vaccines was the most damaging factor. The result is that now, the time it takes to develop a vaccine is increasing. A better balance of basic and applied research could restore a product-oriented focus.

Government-sponsored research of bioweapon countermeasures has generated knowledge and publications, but faltered in delivering practical results. Biodefences cannot be put in place solely by accepting or dismissing research hypotheses; they require safe and effective pharmaceutical products. We need the biological equivalent of the defences that now protect Hawaii's Pearl Harbor.

So what should society and governments do next? At this point, the book's contributions fade. Hoyt dismisses most vaccine stockpiles as unresponsive to today's threats, yet doesn't indicate any need for quick-acting antibody formulations or therapeutics. Instead, she advocates an emphasis on multipurpose technologies and 'platforms', many of which are hypothetical. How can we measure success using a platform approach? How quickly could platforms be transformed into defences against bioweapons?

Perhaps Hoyt's idea is the right interim solution for today's nadir in national will. Perhaps later, if and when it returns, the well-coordinated, product-centric development of vaccines can resume. History offers us lessons in how to do so. ■

---

**John Grabenstein** *is a senior medical director at Merck Vaccines, West Point, Pennsylvania, USA.*
*e-mail: john_grabenstein@merck.com*

---

ILLUSTRATION BY ALESSANDRO GOTTARDO

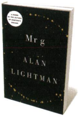
FICTION

# Cosmic creation

**Pedro Ferreira** explores Alan Lightman's latest novel — a magical-reality take on the origins of the Universe.

When a physics heavyweight is mentioned in the same breath as Salman Rushdie and Italo Calvino, it is tough for a reviewer. Few venture into air that rarefied and make it out alive. But when the book is *Mr g*, a creation myth by physicist Alan Lightman, it is worth the risk.

In *Mr g*, Lightman has taken the core of what we know about the origins of the Universe from physics, chemistry and biology and wrapped a few characters around it. The protagonist is the narrator: god, dubbed Mr g. Mr g lives in a timeless Void with his Aunt Penelope and Uncle Deva, playing out skits that could have been lifted from a Woody Allen film — but with the humour on mute. One day, Mr g wakes up from a nap, decides to create a Universe called "Aalam-104729", and from then on marvels at his creation as it evolves and becomes more complex — from the beginning of space and time, to the emergence of fundamental laws, particles, forces, stars, galaxies, planets and, ultimately, life itself.



**Mr g: A Novel About the Creation**
ALAN LIGHTMAN
*Pantheon/Corsair: 2012. 224 pp. $24.95/£9.99*

As sentient beings finally emerge out of the cosmic mess, Mr g is torn: should he intervene or let them go their own way? Throughout, he is taunted by the creepy Belhor, a devilish character (a fine role for Al Pacino if this were ever made into a film) and Belhor's annoying daemons, the Baphomet siblings. Belhor pushes Mr g to allow his creations to do their own thing, and watches with glee as evil and unhappiness begin to emerge — leaving Mr g to observe as, for instance, an impoverished young woman anguishes over stealing meat to feed her starving siblings. Lightman uses the exchanges between Mr g and Belhor to riff on good and evil, free will and relative morality.

Lightman's grasp of the science, in all its gory detail, is unerring. His ability to interweave the fantastical with the factual is impressive — not surprising, given his background. In the 1970s, Lightman established himself as an astrophysical relativist, writing several important papers and books. I still use them. In 1993, he published a magical collection of short stories, *Einstein's Dreams* (Pantheon), in which he took the strange concepts and consequences of Einstein's principle of relativity and wove them into vignettes. I was shocked at how well it worked — and I was not alone. The book has been translated into 30 languages and has led to several stage productions around the world.

Lightman had done something that I had thought impossible: he had brought in the hard science and softened it up. *Einstein's Dreams* is sensual; it breathes. *Mr g* is different. Much more abstract and almost pedagogical, it is a detailed description of the birth and evolution of the Universe that reads like a Rushdiesque fable about an invented place and time. The facts are faultless. Yet the book can occasionally be unintelligible.

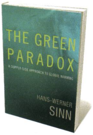
For example, Lightman defines a tick of a clock in terms of a particular frequency of the hydrogen atom. Beautifully precise and clear. But as a result, any other timescales that he mentions must be written in scientific notation. And, of course, other big numbers — such as the number of neurons in a brain or of Universes in what he calls the Void — can also be accurately presented only in scientific notation. You begin to wonder about the readership. Are these nuggets meant to be seen as just icons, like hieroglyphs? Or is the generalist supposed to know what they mean?

It has to be said too that Lightman's work could have worked well as a short story. It loses steam early on and sometimes feels like interspersed vignettes on science and morality with a dose of 'magic' thrown in. Finely crafted it may be, but it can be hard going. Throughout, I had a pretty good idea of what was going to happen and, given the nature of the story, the characters didn't need to be developed. There was nothing to wait for.

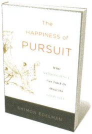
Yet ultimately, this is a marvellous counterpoint to all of the other nonsense out there on creation. Lightman writes exquisitely, so this fable on the origin of space, time, matter and life is a wordfest that is securely pinned to the rational — making him a 'magic realist' of a refreshingly different stripe. ■

**Pedro Ferreira** *is professor of astrophysics at the University of Oxford, UK, and a member of the Oxford Martin School.*
*e-mail: p.ferreira1@physics.ox.ac.uk*

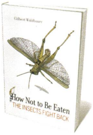
# Books in brief



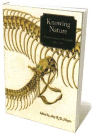
**The Green Paradox: A Supply-Side Approach to Global Warming**
*Hans-Werner Sinn* MIT PRESS *288 pp. $29.95 (2012)*
Frustrated by "counterproductive" climate policies, economist Hans-Werner Sinn proposes a radical alternative. Policy-makers, he notes, have ignored the oil sheikhs and coal barons who supply the fossil-fuel market. Yet it is they who call the shots, as shown by the 'green paradox' — announcements of future reductions in carbon consumption that drive carbon-resource controllers to bump up production. Sinn's antidote to the ideology that plagues policy-making is a "Super-Kyoto" system: unified countries, coordinated caps and trade, and taxation designed to curb the 'extraction habit'.



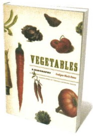
**The Happiness of Pursuit: What Neuroscience Can Teach Us About the Good Life**
*Shimon Edelman* BASIC BOOKS *256 pp. $25.99 (2012)*
The mind, says cognitive psychologist Shimon Edelman, is a literal "meat computer". Our experience of the world is a series of computations carried out by neural wetware. But where does that leave philosophical conundrums such as joy? Taking passages by luminaries including Homer, William Shakespeare and Jorge Luis Borges as touchstones, Edelman powers along on his "quest for an algorithmic understanding of happiness", revealing that it is this computational journey itself that constitutes the good life.



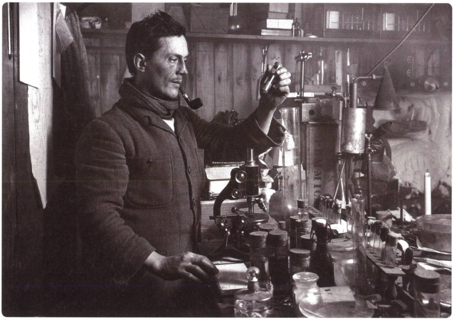
**How Not to Be Eaten: The Insects Fight Back**
*Gilbert Waldbauer* UNIVERSITY OF CALIFORNIA PRESS *240 pp. $27.95 (2012)*
Insects and their predators use a vast, often bizarre array of strategies to eat or to avoid being eaten. Entomologist Gilbert Waldbauer tours tactics on both sides, from sticky lures to kamikaze speed. His fascinating cast ranges from goatsuckers (birds in the order Caprimulgiformes) that can catch massive moths in their bristled gapes, to the brilliantly hued lubber grasshopper (*Romalea guttata*), which, when threatened, both vomits a noxious substance and hisses as stinking froth erupts from its thorax.



**Knowing Nature: Art and Science in Philadelphia, 1740–1840**
*Edited by Amy R. W. Meyers and Lisa L. Ford* YALE UNIVERSITY PRESS *432 pp. $65 (2012)*
When writer Thomas Paine saw Philadelphia in 1774, he noted that nearly every citizen had "some scientific interest or business" — a state that persisted to the mid-nineteenth century. The city boasted artist-naturalists such as John James Audubon, and the study of nature filtered through all strata of society. These 14 illustrated essays, edited by art historians Amy Meyers and Lisa Ford, trace a century of influential scientific endeavour, from botanic gardens and cabinets of curiosity to pioneering colour-plate techniques.



**Vegetables: A Biography**
*Evelyne Bloch-Dano (translated by Teresa Lavender Fagan)* UNIVERSITY OF CHICAGO PRESS *128 pp. $20 (2012)*
Next time you chop a carrot, spare a thought for the centuries of breeding, farming technology and trade that brought it to your kitchen. In brief 'biographies' of 11 vegetables, from pumpkins and artichokes to chard, writer Evelyne Bloch-Dano serves up a feast of associated genetics, agriculture, history and culture. From the role of chillies in world trade to the strange link between peas and the court of Versailles in France, there is much to savour.



Biologist Edward Atkinson in the lab at Cape Evans hut, base camp of the Antarctic Terra Nova expedition.

ANTARCTICA

# Scientists to the end

**Colin Martin** celebrates a London exhibition revealing the research legacy of Robert Scott's final journey.

**Scott's Last Expedition**
*Natural History Museum, London.*
*20 January to 2 September 2012.*

**These Rough Notes**
*Scott Polar Research Institute, Cambridge, UK.*
*Until 5 May 2012.*

It is one of the most poignant moments in polar history. One hundred years ago this week, British Antarctic Expedition leader Robert Falcon Scott and his four companions reached the spot where, 33 days previously, Roald Amundsen and his party had planted a Norwegian flag. All five members of Scott's party died while attempting to return to their base camp at Cape Evans.

That tragic outcome has tended to overshadow the expedition's scientific achievements. In *Scott's Last Expedition*, London's Natural History Museum aims to redress the balance. As curator Elin Simonsson explains, "We want people to know that Scott's British Antarctic Expedition was not simply a quest to reach the South Pole, but an important scientific expedition that carried out work across many fields."

Those fields included glaciology (then in its infancy), biology, magnetism, meteorology, chemistry and physics. The scientific team that participated in Scott's Terra Nova expedition of 1910–13 numbered 12 — the largest sent to the Antarctic at that time — including zoologist Edward Wilson, geologist T. Griffith Taylor and meteorologist George Simpson.

When the ship *Terra Nova* returned to England, it was laden with specimens of 2,109 animals, plants and fossils — with more than 401 new to science. Geological specimens (1,919 of which are in the museum's permanent collection) helped to construct knowledge of the continent. The meteorological data — the longest unbroken record of weather collected in the early twentieth century — provide baselines that are now used in assessing climate change.

➲ **NATURE.COM**
For more on Scott's legacy:
go.nature.com/9fbuuj

The shore party of 25 men lived and worked in a wooden hut measuring 112 square metres (about half the area of a singles tennis court). Although the hut was equipped with scientific instruments, the scientists' real lab was the Antarctic. The exhibition includes a full-size representation of the hut (the original is preserved at Cape Evans), with actual artefacts from the expedition on display, including scientific apparatus and notebooks as well as a range of biological and geological specimens.

These include the shells of three eggs from the emperor penguin (*Aptenodytes forsteri*; pictured), which contained embryos at different stages of development. Wilson, his assistant Apsley Cherry-Garrard and marine lieutenant Henry Bowers collected them under unimaginably harsh conditions, trudging to the penguin breeding colony at Cape Crozier — a trip immortalized by Cherry-Garrard in his 1922 book *The Worst Journey in the World*. The team's painstaking preservation and study of the embryos ultimately came to little, as theories about bird evolution from reptiles had changed by the time their report was published in 1934.

However, the fossils of the extinct plant *Glossopteris indica*, found by Scott and his final companions, provided important evidence that the climate of Antarctica once supported vegetation and was a part of the supercontinent Gondwana.

The expedition is also documented by photographs, many by photographer Herbert Ponting. These include an iconic image of Scott writing in his diary, and parasitic-infection specialist Edward Atkinson working at the bench. Similarly informative are Scott's own 'lost' photographs. Only a handful of the 120 images he took under Ponting's tutelage had been available until last year, when Wilson's great-nephew, polar historian David M. Wilson, published *The Lost Photographs of Captain Scott* (Little, Brown; 2011). The snowy panoramas and shots of expedition life help to reveal the daily realities of his last endeavour, while polar scientists still use the images for comparison with contemporary photographs to assess climate impacts.

A concurrent exhibition, *These Rough Notes*, is showing at the Scott Polar Research Institute in Cambridge, UK. It takes its title from a journal entry written by Scott shortly before his death — "These rough notes and our dead bodies must tell the tale..." — and displays personal papers of Scott and other expedition members. They record the heartbreaking poignancy of the polar party's death. But as the exhibition at the Natural History Museum shows, we need to reinforce the dual importance of this extraordinary expedition. "We want the scientific work to make the bagging of the Pole merely an item in the results," wrote Wilson at the expedition's outset. He didn't, of course, foresee that his own fate, and that of Scott and other colleagues, would long eclipse an enduring scientific legacy that led to the publication of 81 papers. ■

**Colin Martin** *is a writer based in London.*
*e-mail: cmpubrel@aol.com*

CANTERBURY MUS. NEW ZEALAND
M. WATSON/ARDEA.COM

# Correspondence

## Rare day to highlight rare diseases

In the last leap year in 2008, 29 February was chosen to mark Rare Disease Day by EURORDIS, the European Organisation for Rare Diseases. This year, this rare day will again serve to raise awareness of the plight of patients with rare disorders.

Rare diseases fail to attract commercial research efforts because of the small number of people affected (fewer than 5 in 10,000). Examples include the nerve disorder Charcot–Marie–Tooth disease and Proteus syndrome, which causes excess overgrowth of tissues.

As a result, only a handful of scientists, often working in isolation, dedicate their research to a particular rare disease. But such select studies can deliver groundbreaking insights into more general disease processes. Exposure of molecular mechanisms underlying rare diseases therefore has the potential to help many other patients.

Also on 29 February, the first international congress dedicated to rare diseases will be held in Basel, Switzerland (www.react-congress.org). It will highlight the recently established International Rare Diseases Research Consortium (IRDiRC), which plans to coordinate international policy-making with results from national research projects. These initiatives should eventually translate scientific advances into benefits for patients.

**Susan M. Gasser** *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland.*
*susan.gasser@fmi.ch*
**James R. Lupski** *Baylor College of Medicine and Texas Children's Hospital, Houston, Texas, USA.*
**Yann Le Cam** *European Organisation for Rare Diseases, Paris, France.*
**Olivier Menzel** *BLACKSWAN Foundation, Swiss Foundation for Research on Orphan Diseases, Porza, Switzerland.*

## Asian medicine: small species at risk

The demands of traditional Asian medicine (TAM) don't just pose a threat to the survival of tigers and rhinos (*Nature* **480,** S101–S103; 2011). Numerous smaller species are also at risk, as a result of being traded in large volumes.

For example, millions of dried seahorses (*Hippocampus* spp.) and Tokay geckos (*Gekko gecko*) are sold annually for use against impotence and circulatory problems in TAM. Gecko sales have been further fuelled by rumours that larger individuals can cure HIV. Trade in dried orchid parts for eye ailments and in ginseng rhizomes is consuming plants in their millions. The decline in Asia's turtle population is being aggravated by the marketing of TAMs that contain turtle plastrons (the flat underbelly of the shell) to treat skin diseases.

The Convention on International Trade in Endangered Species seeks to protect most of these globally threatened species. The high demand for scaly anteaters, or pangolins (*Manis* spp.), for example, has led to an international trade ban. However, the scales of tens of thousands of pangolins are still confiscated annually in Asia.

For some species, captive breeding and plant propagation have been successful, but the demand for wild-sourced ingredients means that such techniques cannot wholly mitigate the impact of trade.

The killing of endangered tigers and rhinos for TAM will not go unnoticed. But the insidious destruction of smaller species must also be halted by enforcing stricter regulations.

**Vincent Nijman, K. Anne-Isole Nekaris** *Oxford Brookes University, Oxford, UK.*
*vnijman@brookes.ac.uk*
**David P. Bickford** *National University of Singapore, Singapore.*

## Call to split fisheries at home and abroad

I agree that "no one should doubt that our seas need protection" (*Nature* **480,** 151; 2011) and that establishing marine protected areas alone will not do the job (*Nature* **480,** 14–15; 2011). A combination of measures is needed, including the elimination of overfishing subsidies. These incentives were introduced when fisheries seemed inexhaustible, but they inflate profitability and drive fishing beyond economic or sustainable levels.

The global community mandated the World Trade Organization (WTO) to discipline overfishing subsidies more than ten years ago, but the issues were still unresolved at their meeting last month. One reason is that WTO negotiators are trying to broker an all-inclusive deal that encompasses domestic and international, small- and large-scale fisheries. But this approach is hindered by national interests.

The answer is to split the world's fisheries into domestic and international ones. Domestic fisheries would operate within a country's economic exclusion zone and target fish stocks that spend all their lives there. This split is necessary because the incentives to eliminate overfishing subsidies differ according to whether a fishery is domestic or international.

For a domestic fishery, the heavy lifting should be on the home front; for an international fishery, global coordination would be needed, because unilateral action by one country will not eliminate overfishing. Categorizing fisheries in this way would make it easier to identify leverage points for eliminating overfishing subsidies.

**U. Rashid Sumaila** *University of British Columbia, Vancouver, British Columbia, Canada.*
*r.sumaila@fisheries.ubc.ca*

## Investors unfazed by drug-patent expiry

Your data on changing stock prices for five pharmaceutical companies from 1997 to 2010 (*Nature* **480,** 16–17; 2011) erroneously indicate a 39% drop in aggregated share value, when in fact it would have risen by 82% (see Correction, *Nature* **480,** 425; 2011). The error was due mainly to the selection of unadjusted, rather than adjusted, stock-market closing prices.

The corrected data set indicates that the aggregated share prices rose roughly in line with the Dow Jones Industry Average. Considering that these companies have the biggest patent cliffs in the industry, this result is very positive; moreover, it argues against your implication that the impending expiry of their patents is affecting investor confidence in pharmaceutical stocks.

Investors accept that the strength of a company's patent portfolio is only one of several factors. Hence, despite the impending expiry of patents for blockbuster products, including Pfizer's Lipitor, the performance of the five worst-hit companies was comparable with that of other sectors for the period 1997–2010.

Other factors such as decreasing revenues can also affect investor confidence. Your simplistic view that the patent cliff is the only contributor does an injustice to the efforts of pharmaceutical companies and governments, and unnecessarily perpetuates a false uncertainty.

**David Brindley** *University College London, UK; Harvard Stem Cell Institute and Harvard Business School, Massachusetts, USA.*
*david_brindley@harvard.edu*
**Brock Reeve** *The Harvard Stem Cell Institute, Cambridge, Massachusetts, USA.*
**Chris Mason** *University College London, UK.*

# Paul Mead Doty

## *(1920–2011)*

### Chemist and arms-control advocate who unravelled the structures of biomolecules.

Paul Mead Doty, who died of heart failure at home on 5 December 2011 at the age of 91, leaves three legacies: his revelation of the structures of numerous proteins and nucleic acids; his contribution to our very survival through his tireless activity on behalf of arms control and disarmament; and the many successful scientists and policy figures whom he mentored.

Doty was born on 1 June 1920 in Charleston, West Virginia, and was schooled in Pennsylvania. His exceptional talent was recognized at high school, where a chemistry teacher asked him to take over the class. Doty earned a Bachelor's degree in chemistry at Pennsylvania State University (then Pennsylvania State College) in University Park in 1941 and a PhD in chemistry at Columbia University in New York in 1944. He was renowned for being late. On the first day of classes at Columbia, as the students were seated in alphabetical order, he missed his place. So Doty sat at the end, and thereby made the acquaintance of Bruno Zimm, a friendship that lasted a lifetime and would lead to seminal scientific collaborations.

During his doctorate, Doty measured the electron affinity of bromine. He moved into the new field of polymer science and, with Bruno Zimm, pioneered the theory and use of light scattering to characterize polymer molecules in solution. A postdoctoral position at the Polytechnic Institute of Brooklyn, New York, brought Doty into contact with founders of polymer science, such as Herman Mark, and alerted him to the unusual chemistry and physics of macromolecules. During a second postdoc, at the University of Cambridge, UK, Doty learned, mainly from Max Perutz, about the X-ray diffraction of protein crystals, which allowed their structures to be determined.

#### BIOLOGICAL MACROMOLECULES

After briefly joining the faculty at Notre Dame University, Indiana, Doty accepted an assistant professorship at Harvard University in Cambridge, Massachusetts, in 1948. He remained at Harvard for the rest of his life. During his tenure there, he founded the Department of Biochemistry and Molecular Biology and the Center for Science and International Affairs.

Doty's research focus shifted at Harvard to large biological molecules. Realizing that proteins could not be random polymers, because they could be crystallized and had unique structures, he established a programme to study synthetic polypeptides. These chains of amino acids allow structural transitions, such as from helix to random coil, to be followed by a technique called optical rotatory dispersion. By tracing structural transitions induced by changes in temperature, he was able to estimate the fraction of each protein that was helical. These results laid the foundations for modern studies of protein folding.

S. WILKE/HKS BELFER CENTRE

He then, in the 1950s, turned to nucleic acids, a step made possible by the availability of synthetic polynucleotides. Out of this work came the discovery that DNA structure could be altered using heat, that the melting temperature of DNA from various bacteria differed according to the fraction of G and C nucleotides in it, and that RNA contains regions of secondary structure. These observations are the basis of much modern biotechnology, including DNA sequencing and the polymerase chain reaction.

Doty's strength lay in identifying important scientific issues with conceptual and practical implications that could be resolved by devising simple model systems. He was adept at finding the right person for a particular job, hiring Jacques Fresco for the polynucleotide work, for example, and Julius Marmur for studies using DNA. He was discriminating in how he worked with his students, balancing a laissez-faire attitude with specific guidance according to each student's needs. As PhD students, we both benefited from this style of mentorship.

#### PUBLIC AFFAIRS

His most important choice of associate was a graduate student, Helga Boedtker, whose thesis on the structure of collagen in solution became a landmark. They married, and Helga became the mainstay of the laboratory for more than 40 years until her death in 2001. When Doty was absent from the lab during his trips abroad to promote arms control, a quiet word to Helga brought urgent matters to his attention.

Very early in his career, Doty became active in science and public affairs, with a focus on US–Soviet relations. He chaired the Federation of American Scientists and the US Pugwash Committee, worked on disarmament plans under President Dwight Eisenhower's science adviser James Killian, and was a member of President John F. Kennedy's Science Advisory Committee. Doty co-chaired the Soviet–US disarmament committee (now the Committee on International Security and Arms Control of the National Academy of Sciences), helped to arrange the first visits to Moscow of US secretaries of state Henry Kissinger and Cyrus Vance and played a key part in developing the basis for the US–Soviet Anti-Ballistic Missile Treaty.

As noted by molecular biologist Matthew Meselson, Doty "surely helped to avoid catastrophe". ■

---

**Stuart A. Rice** *is in the Department of Chemistry and the James Franck Institute, University of Chicago, Illinois 60637, USA.*
**Robert Haselkorn** *is in the departments of chemistry, and of molecular genetics and cell biology, University of Chicago, Illinois 60637, USA.*
*e-mail: rh01@uchicago.edu*

# RESEARCH

## NEW ONLINE

*Papers published this week at Nature.com/nature*

**NEWS & VIEWS** The role of nano-particle shape in self-assembly of new materials **p.268**

**REVIEW** The Amazon basin may be changing from carbon sink to carbon source **p.321**

**LETTER** How foaming magmas bring diamond-containing minerals to the surface **p.352**

**15 JANUARY 2012**

**PLANT SCIENCES**

### Where and when auxins function

Auxins are key factors in the regulation of plant growth and development, but effective tools to analyse their distribution and signalling are in short supply. Teva Vernoux and colleagues have developed a sensor engineered to monitor auxin responses and distribution in tissues. Based on an auxin/indole acetic acid protein, DII-VENUS provides high-resolution spatio-temporal information about auxin distribution and response during plant development.

**A novel sensor to map auxin response and distribution at high spatio-temporal resolution**
G Brunoud *et al.*
doi:10.1038/nature10791

**IMMUNOLOGY**

### Sensing dangerous bacteria

The galectins are carbohydrate-binding proteins with a range of functions inside and outside the cell. They accumulate in the cytosol, which is normally devoid of complex carbohydrates, marking them out as prime candidates for a role as danger or pattern-recognition receptors. Teresa Thurston *et al.* now identify galectin 8 as a danger receptor, protecting cells against bacterial infection. It binds to host glycans exposed on damaged bacteria-containing vesicles and recruits the ubiquitin-binding autophagy receptor NDP52 to clear the cytosol of invading bacteria.

**Galectin 8 targets damaged vesicles for autophagy to defend cells against bacterial invasion**
T L M Thurston *et al.*
doi:10.1038/nature10744

**METHODS**

**Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism**
M L Steinhauser *et al.*
doi:10.1038/nature10734

**CELL BIOLOGY**

**Multi-isotope imaging mass spectrometry reveals slow protein turnover in hair-cell stereocilia**
D-S Zhang *et al.*
doi:10.1038/nature10745

**18 JANUARY 2012**

**NEUROSCIENCE**

### Brain's response to reward

The ventral tegmental area (VTA), a part of the brain involved in reward and addiction, contains both dopaminergic and GABAergic signals, but the role of the subpopulations in reward representation and processing is unclear. Using optogenetics to identify dopaminergic and GABAergic neurons, Jeremiah Cohen *et al.* characterize VTA responses to rewarding and aversive stimuli. The two neuronal types have distinct responses to reward: dopaminergic neurons signal reward prediction errors; GABAergic neurons signal reward expectation.

**Neuron-type-specific signals for reward and punishment in the ventral tegmental area**
J Y Cohen *et al.*
doi:10.1038/nature10754

**CELL BIOLOGY**

### Autophagy boosted by exercise

Exercise has beneficial effects on human health, including protection against metabolic disorders such as diabetes. The cellular mechanisms underlying these effects are unclear, but one possibility is a stimulatory effect on autophagy, which acts to release nutrients into the cell through the destruction of unwanted proteins. It also protects against various diseases in animal models. Beth Levine and colleagues now show that acute exercise induces autophagy in skeletal and cardiac muscle of fed mice. Mice that are incapable of exercise-induced autophagy have decreased endurance and altered glucose metabolism during acute exercise, as well as impaired chronic exercise-mediated protection against high-fat, diet-induced glucose intolerance.

**Exercise-induced BCL2-regulated autophagy is required for muscle glucose homeostasis**
C He *et al.*
doi:10.1038/nature10758

**PSYCHOLOGY**

### Ageing, genetics and cognition

The world's population is ageing at an unprecedented rate, adding urgency to the study of the causes of mental deterioration with age. The Scottish Mental Surveys of 1932 and 1947 subjected 11-year-old children to a well-validated test of general intelligence. Almost 2,000 of these individuals have been tested again since — and had genome-wide single nucleotide polymorphism data collected — to provide a unique data set of lifetime cognitive measures. The data suggest that genetic factors underlie the differences in intelligence stability across the human lifespan. Estimates of the genetic contributions to cognitive stability and change in this cohort imply that a search for the genetic causes of lifetime cognitive change is worthwhile.

**Genetic contributions to stability and change in intelligence from childhood to old age**
I J Deary *et al.*
doi:10.1038/nature10781

**MATERIALS SCIENCE**

**Pathway complexity in supramolecular polymerization**
P A Korevaar *et al.*
doi:10.1038/nature10720

**BIOCHEMISTRY**

**Genome-wide structure and organization of eukaryotic pre-initiation complexes**
H S Rhee & B F Pugh
doi:10.1038/nature10799

**CANCER**

**DNA breaks and chromosome pulverization from errors in mitosis**
K Crasta *et al.*
doi:10.1038/nature10802

## ➲ NATURE PODCAST

Eric Davidson, first author of the Review on the state of the Amazon basin (page 321), discusses the effects of human activities in the region. And Cesar Harada talks about how crowd-sourced funding is helping to support the Protei project, which uses pollution-collecting sailing drones to clean up oil spills. **nature.com/nature/podcast**

# NEWS & VIEWS

MATERIALS SCIENCE

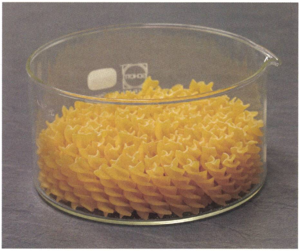
# A fresh twist for self-assembly

**Molecular helicity affects many of the bulk properties of materials. A study finds that helicity also controls the self-assembly of colloidal particles, opening the door to a new generation of functional materials. SEE LETTER P.348**

VOLKER SCHALLER
& ANDREAS R. BAUSCH

The next time you go to the supermarket, take a look at the pasta. You'll probably find everything from long, thin spaghetti to butterfly-shaped farfalle and twisted fusilli. On closer inspection, you'll see that the strands of spaghetti readily align and pack closely together, whereas the packing of the fusilli is considerably more complex. This complexity is due to the fusilli's chirality — its helical geometry. On page 348 of this issue, Gibaud *et al.*[1] report that such complexity of packing can be exploited to control the self-assembly of nanometre-scale particles, allowing the reversible formation of various architectures*.

To closely pack two individual pieces of fusilli, the pasta pieces have to twist with respect to each other so that their long axes are not parallel — this is an effect of their chirality. But when many fusilli are crammed together, this twist hinders the ability of the pasta to align in one direction, as required for efficient packing. The resulting competition between chirality and packing can have astonishing consequences for geometrically constrained chiral objects: if you stack fusilli upright in a beaker, the axes of the fusilli in the centre are perfectly vertical, whereas those at the edges of the beaker twist away from this alignment (Fig. 1).



V. SCHALLER

**Figure 1 | Pasta packing.** When constrained in a circular container, fusilli pasta pieces mostly pack together so that their long axes are vertically aligned. But at the edges of the container, the pasta twists away from this alignment. This packing arrangement is a consequence of the pasta's helicity (chirality). Gibaud *et al.*[1] report that the chirality of colloidal particles affects the self-assembly of those particles.

This balance between geometric constraints and chiral interactions is at the heart of the unique properties of chiral liquid crystals. In 'blue' liquid-crystal phases, for instance, chiral molecules self-assemble into cylinders of twisted molecules that stack in a cubic lattice, giving rise to vividly coloured and reflective materials[2]. And in smectic liquid crystals, chiral interactions lead to a twisted, layered structure that is also manifested in the optical and mechanical properties of the materials[3].

Gibaud *et al.*[1] examined the balance between chirality and geometric constraints in the self-assembly of colloidal particles. Their model system was an ensemble of rod-like fd viruses, 7 nanometres in diameter and 880 nm in length, suspended in water. The chirality of these viruses depends on the ambient temperature: the rods are achiral at high temperatures, but chiral at low temperatures. The authors observed that, as for any other suspension of colloidal particles[4], the addition of an appropriate polymer that does not adsorb to the rods drove a phase separation of the mixture into a polymer-rich and a colloid-rich phase. At high temperatures, at which the rods are achiral, this led to the self-assembly of a circular colloidal membrane (a monolayer of vertically aligned rods[5]).

The authors observed that the rods at the centre of the membrane readily align parallel to each other, but at the perimeter of the membrane they twist gradually away from the vertical. This twist at the margin minimizes the interface between the polymer-rich and the rod-rich phases, and so reduces the interfacial tension that arises from the imbalance of forces between particles in this region. But twisting rods out of alignment with the other rods in the membrane has its own energetic cost, known as elastic energy. The formation of circular membranes with twisted margins is therefore the result of a trade-off between minimizing the phase interface and minimizing the elastic energy.

But what happens if the rods are chiral, so that twisted packing is preferred — just as it is for closely packed fusilli? Gibaud *et al.* addressed this question by performing experiments at lower temperatures, thereby 'switching on' the chirality of the viruses. They observed that increases in chirality — that is, increases in the contribution of chiral interactions to the energy balance of the system — reduce the elastic energy, thus lowering the energetic cost of creating a twist at the membrane's margin. This destabilizes the edges of the circular membrane and triggers the formation of ribbon-like structures that splay out from the circular membrane (see Fig. 3a of the paper[1]). The effect of increasing the chirality of the viruses in the colloidal membrane is therefore similar to that of adding surfactants to oil–water mixtures: it decreases the interfacial tension and so allows more interfaces to form.

The beauty of the authors' experimental set-up is that it allows all the parameters involved in the self-assembly of chiral colloids to be readily teased out and measured. Another advantage of the system is that the colloidal building blocks are much bigger than the molecules typically used in studies of chiral materials. This allowed Gibaud *et al.* to precisely quantify the colloidal self-assembly process at all scales, from the microscopic movements of individual viruses to the macroscopic properties of the resulting membrane. It also allowed them to exert well-defined local forces on the system, to affect the balance between twist and interfacial energy, and so to induce structural transitions — for example, by pulling on the membrane, they converted it into ribbon structures.

By including chiral interactions in their system, Gibaud *et al.* have added a new degree of freedom to the self-assembly of colloidal particles. Transferring their results to other

---

*This article and the paper[1] under discussion were published online on 4 January 2012.

materials opens the door to the hierarchical assembly of functional materials that could react sensitively to changes in ambient conditions or to mechanical stimuli. One way to extend the functionality of their system could be to make colloidal particles that are not only chiral, but also have surfaces patterned with microscopic domains that attract or repel each other. Such patterned colloidal particles have recently been shown to self-assemble into useful functional structures such as 'kagome' lattices[6].

Chiral interactions between particles occur in all kinds of materials, from liquid crystals[2,3] to cytoskeletal filaments in cells (for which the hierarchical assembly of bundle-like structures is dependent on, and sensitive to, the helical twist of the filaments[7,8]). Most of the chiral materials studied so far are passive or in thermal equilibrium. This means that their structural assembly is governed only by diffusion and by local interactions between the constituent particles. But most naturally occurring materials are far from passive. Instead, they constantly consume energy so that their particles self-organize into higher-order structures — the cytoskeletons of cells are prime examples of this[9,10].

The next step, therefore, is to apply the principles identified by Gibaud and colleagues[1] to active systems. Imagine replacing the viruses used in this study with self-propelling chiral bacteria — would they still pack together in a two-dimensional membrane, and, if so, would the membrane move, or rotate, across mesoscopic or macroscopic distances? Could actively beating ribbons form, or would the noise stemming from active movement of the bacteria prevent the formation of such structures? All we can say for certain is that chiral interactions will add a new twist not only to the self-assembly of colloids, but also to the self-organization of active materials. ■

**Volker Schaller** *and* **Andreas R. Bausch** *are at the Lehrstuhl für Biophysik-E27, Technische Universität München, Garching 85748, Germany.*
*e-mail: abausch@mytum.de*

1. Gibaud, T. *et al. Nature* **481,** 348–351 (2012).
2. Coles, H. J. & Pivnenko, M. N. *Nature* **436,** 997–1000 (2005).
3. Hough, L. E. *et al. Science* **325,** 456–460 (2009).
4. Aarts, D. G. A. L., Schmidt, M. & Lekkerkerker, H. N. W. *Science* **304,** 847–850 (2004).
5. Barry, E. & Dogic, Z. *Proc. Natl Acad. Sci. USA* **107,** 10348–10353 (2010).
6. Chen, Q., Bae, S. C. & Granick, S. *Nature* **469,** 381–384 (2011).
7. Claessens, M. M. A. E., Semmrich, C., Ramos, L. & Bausch, A. R. *Proc. Natl Acad. Sci. USA* **105,** 8819–8822 (2008).
8. Grason, G. M. & Bruinsma, R. F. *Phys. Rev. Lett.* **99,** 098101 (2007).
9. Kasza, K. E. *et al. Curr. Opin. Cell Biol.* **19,** 101–107 (2007).
10. Köhler, S., Schaller, V. & Bausch, A. R. *Nature Mater.* **10,** 462–468 (2011).

---

GENOMICS

# The path to retinoblastoma

**Genomic analyses of tumours of the childhood cancer retinoblastoma reveal a low mutation rate, challenging the view that genomic instability is crucial for its progression. The work also identifies a new therapeutic target.** SEE ARTICLE P.329

**JULIEN SAGE & MICHAEL L. CLEARY**

Retinoblastoma is a rare tumour that affects retinal cells in the eyes of children. Analyses of familial and sporadic cases of this cancer, backed by studies in genetically engineered mice, have shown that loss of function of the tumour-suppressor protein RB1 (also known as RB) is required for the development of most, if not all, tumours of this type. However, it is not clear how RB1-deficient retinal cells progress to malignant tumour cells[1]. In addition, emerging evidence that loss of RB1 function can induce genomic instability[2] has raised the tantalizing possibility that RB1-deficient retinal cells might be predisposed to accumulating many additional mutations, further complicating the identification of mutations that contribute to the development and maintenance of retinoblastoma. On page 329 of this issue, however, Zhang *et al.*[3] demonstrate that retinoblastoma genomes have very few recurrent mutations in genes other than *RB1*. Instead, the expression of cancer-related genes is affected by epigenetic modifications on chromosomes, which do not affect DNA sequence but are inherited after cell division.

To identify mutations that could cooperate with loss of RB1 function in tumour development, Zhang and colleagues[3] sequenced and compared the genomes of normal tissue and retinoblastoma tumours from four patients. The researchers found that *RB1* was the only known cancer-related gene consistently mutated, and that the retinoblastomas had 15-fold fewer total mutations than other types of solid tumour whose genomes have been sequenced[4].

Next, the authors searched for epigenetic alterations and for abnormal gene expression in retinoblastoma cells. They identified the gene that encodes the protein kinase enzyme SYK as a potential oncogene whose expression is consistently higher in retinoblastoma cells than in normal immature retinal cells. Moreover, the activity of SYK was essential for the growth of retinoblastoma cells. The authors also show that certain small molecules that selectively inhibit SYK activity induce the death of retinoblastoma cells in a mouse model of the disease.

These findings[3] indicate that SYK may be a promising target for treating patients with retinoblastoma. *SYK* was not identified in previous searches for genes with a role in this cancer because it is not mutated or structurally rearranged in retinoblastoma. Therefore, Zhang and colleagues' study emphasizes the importance of high-throughput approaches that integrate genome sequencing with gene-expression analysis and epigenomics to identify cancer genes.

Low mutation frequency has also been observed[5] in medulloblastoma (a type of brain tumour that affects children), suggesting that it could be a general feature of childhood cancers. A possible explanation for this difference between the cancers of children and adults is that childhood tumours arise in cells that are naturally undergoing rapid developmental growth, with fewer brakes on their proliferation than cells in adults. An alternative explanation is that, in children, these cells are negotiating crucial developmental checkpoints that are susceptible to corruption, leading to incomplete or abnormal maturation. In both cases, only a few mutations would be needed to trigger the cellular changes associated with cancer.

Furthermore, epigenetic changes in children and excessive mutations in adults may have similar roles in cancer development. Another childhood cancer, Wilms' tumour, also has a relatively stable genome and displays an increased variation in DNA-methylation patterns compared with normal cells[6]. RB1-deficient retinal cells may be particularly susceptible to this tumour-formation mechanism, because RB1 interacts with the machinery that controls the epigenetic status of cells, including enzymes that remodel chromatin (DNA–protein complexes) and other enzymes that add methyl groups to DNA[1]. Thus, childhood cancers such as retinoblastoma may carry epigenetic abnormalities that change gene expression and are stably propagated through subsequent cell divisions, helping to maintain tumour-associated features.

If this model is correct, it is possible that RB1-deficient tumours in adults — such as small-cell lung cancer and some breast, prostate or blood cancers — also display epigenetic changes that are crucial for tumour development and maintenance. Future experiments, including integrated genomic analyses, may provide the answer.

The underlying mechanism of alterations in gene expression by epigenetic means in retinoblastoma is still unclear. But Zhang and co-workers' observation[3] that 13% of these tumours show recurrent mutations in the *BCOR* gene offers a possible explanation. The BCOR protein is highly expressed in the fetal retina and is essential for eye development, as evidenced by a congenital eye disorder, syndromic microphthalmia, which results from a heritable mutation in *BCOR*[7]. BCOR associates with proteins that repress gene expression epigenetically[8]. This raises the possibility that loss of BCOR function, due to acquired mutations in its gene, may contribute to an altered epigenetic landscape in RB1-deficient retinal cells. Consequently, inappropriate expression of crucial genes may impair cell maturation and so facilitate the progression of retinoblastoma. It will be of interest to determine whether BCOR is part of the repression machinery that silences SYK expression in the normal retina, and whether acquired *BCOR* mutations provide at least one route to altered expression of SYK in retinoblastoma.

Nevertheless, the current work — using a comprehensive, integrated genomics approach — is notable not just for demonstrating that epigenetic alterations have a predominant role in the progression of retinoblastoma. The new possibilities it raises for therapies in this childhood malignancy, and possibly in other types of tumour, are equally noteworthy. ■

**Julien Sage** *is in the Departments of Pediatrics and of Genetics and* **Michael L. Cleary** *is in the Departments of Pathology and of Pediatrics, Stanford Cancer Institute, Stanford University School of Medicine, Stanford, California 94305, USA.*
*e-mails: julsage@stanford.edu; michael.cleary@stanford.edu*

1. Chinnam, M. & Goodrich, D. W. *Curr. Top. Dev. Biol.* **94,** 129–169 (2011).
2. Manning, A. L. & Dyson, N. J. *Trends Cell Biol.* **21,** 433–441 (2011).
3. Zhang, J. *et al. Nature* **481,** 329–334 (2012).
4. Stratton, M. R. *Science* **331,** 1553–1558 (2011).
5. Parsons, D. W. *et al. Science* **331,** 435–439 (2011).
6. Hansen, K. D. *et al. Nature Genet.* **43,** 768–775 (2011).
7. Ng, D. *et al. Nature Genet.* **36,** 411–416 (2004).
8. Fan, Z. *et al. Nature Cell Biol.* **11,** 1002–1009 (2009).

---

EVOLUTIONARY BIOLOGY

# A ratchet for protein complexity

**Molecular machines containing related protein subunits are common in cells. Reconstruction of ancient proteins suggests that this type of complexity can evolve in the absence of any initial selective advantage.** SEE LETTER P.360

W. FORD DOOLITTLE

Organisms and cells are bewilderingly complicated, and the molecular machines that perform many basic cellular functions are often giant, multisubunit, multifunctional protein complexes with tangled evolutionary histories. It is generally assumed that such complexes arose by the stepwise accretion of individual proteins, each addition representing a selective advantage by adding to or refining the machine's performance. But on page 360 of this issue, Thornton and colleagues[1] argue against this standard explanation in one particular instance — that of a ring-shaped protein complex in fungi. The authors show how evolutionary processes entailing loss of function rather than gain might initially drive a system towards complexity, independently of selection*.

Many cellular molecular machines contain several proteins that self-assemble into a multi-subunit ring. In simple cases, rings are homo-oligomeric; that is, all the subunits are identical and thus probably the products of a single gene. In more complex (hetero-oligomeric) examples, the protein molecules in the ring are different, but may be related. Often, hetero-oligomeric rings seem to have arisen from homo-oligomeric complexes after a gene encoding a single subunit became duplicated, producing two or more gene copies (called paralogues), with each copy subsequently evolving to encode a slightly different protein.

The question that Thornton and colleagues address is, why does this happen? Do the structurally distinguishable subunits and/or their specific pattern of assembly confer improved or additional functions on the protein complex, with selection for their enhanced performance being the evolutionary driving force? Alternatively, might neutral processes be responsible, at least initially?

To understand how the latter possibility might come about, imagine that a gene encoding, for example, eight subunits of a homo-oligomeric ring undergoes duplication (Fig. 1). The two paralogous gene copies that this duplication produces — and the protein subunits they encode (call them A and B) — will of necessity begin to diverge through accumulation of neutral mutations. These mutations might not at first affect any of the subunits' functions, so that functional octameric rings could continue to form by a random sampling

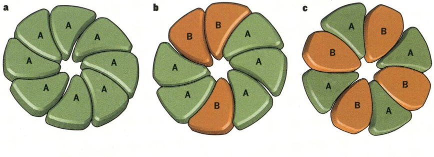


**Figure 1 | Evolution of complexity in a protein ring. a,** Homo-oligomeric protein complexes contain identical subunits (A) that are typically encoded by a single gene. **b,** If the gene is duplicated, the two gene copies can diverge through the accumulation of neutral mutations, generating structurally distinguishable but functionally unaltered subunits A and B, which can form functional hetero-oligomeric rings by random mixture. **c,** If additional mutations prevent subunits from binding to others of their own type, functional rings could still be formed by alternating subunit types. As further mutations accumulate, the probability of returning to the initial homo-oligomeric situation becomes very small. Thus, the subunit composition of a protein association may become complex in the absence of initial selective advantages. Thornton and colleagues[1] provide experimental evidence suggesting that this type of process has occurred in the evolution of a six-membered protein ring that forms part of the vacuolar $H^+$-ATPase enzymes in fungi.

---

*This article and the paper[1] under discussion were published online on 8 January 2012.

of A and B subunits. But further mutations might result in each subunit losing the ability to bind others of its own type, A to A or B to B. However, as long as A still binds B and B still binds A, functional rings can form.

As A and B continue to co-evolve through the acquisition of further neutral mutations, the likelihood of reversing back along the mutational path to the initial single-gene homo-oligomeric complex becomes vanishingly small. We can thus see the homo-oligomer to hetero-oligomer transition as an inexorable evolutionary ratchet. And even if there were some slight selective disadvantage to using a hetero-oligomeric complex (for example, complex formation could become slower), this could be overcome in small populations — in which the efficiency of selection is reduced and mildly harmful mutations can be fixed by chance[2].

Some chaperonins — molecular machines that assist other proteins to fold into functional configurations — have until recently provided the best evidence for such a scenario. Microorganisms known as archaea have one, two or three paralogous genes encoding chaperonin subunits that form rings: homo-oligomeric eight-membered rings when there is a single gene, or hetero-oligomeric eight- or nine-membered rings when there are two or three genes. Such paralogy seems to reoccur frequently in archaea, and a neutral co-evolutionary picture of the sort described above has been suggested[3] to explain it.

Comparative bioinformatic studies[4] support the idea that hetero-oligomeric archaeal chaperonins consisting of two different subunits are not functionally distinct from the homo-oligomers, and thus did arise neutrally. Adaptive changes can follow later, however. Indeed, there is direct evidence for specialized roles for the eight distinct chaperonin subunits of eukaryotes (organisms such as plants, animals and fungi)[5].

So a more realistic picture might be as follows. Rarely, gene duplication is quickly followed by functional differentiation (as proposed in early formulations of the hypothesis of evolution by gene duplication[6]). Selection against loss of either duplicate would then be ensured early on. Perhaps more often, though, the paralogous condition might first be locked in by the above-described co-evolutionary ratchet, allowing functional differences (if any) to be acquired later (if ever). That is what Thornton and colleagues[1] argue has happened in the case of a six-membered hetero-oligomeric ring that forms part of the enzymes known as vacuolar $H^+$-ATPases (V-ATPases) in fungi. Their argument is especially compelling because it rests not only on bioinformatics, but also on experiments, including the synthesis and testing of proteins that the authors propose are ancestral to the V-ATPase subunits.

In most eukaryotes, the V-ATPase six-membered ring is made up of five copies of the protein Vma3 and one of its paralogue Vma16. In fungi such as budding yeast, by contrast, one of the five Vma3 proteins is replaced by Vma11, which is a fungus-specific paralogue of Vma3. The abilities of these subunits to bind to one another[7] are such that Vma11 must lie between Vma16 and Vma3 in the ring, and Vma3 cannot form the interface with the side of Vma16 that Vma11 can. This represents a sort of functional degeneration, because the last common ancestor of Vma3 and Vma11 must have been able to interface with either side of Vma16, as Vma3 must still do in organisms other than fungi.

Thornton and colleagues[1] synthesized proteins that matched the phylogenetically inferred ancestor of Vma3 and Vma11, as well as that of Vma16 and other intermediates in the predicted evolutionary pathway. They demonstrate that Anc.3-11, the inferred ancestor of fungal Vma3 and Vma11, does indeed interface with either side of Vma16. The authors tested the functionality of these and other inferred evolutionary ancestors and intermediates by expressing them in yeast mutants that lacked current versions of one or more of the three genes. According to this picture, when Anc.3-11 underwent duplication, one paralogue (which was to become modern fungal Vma3) lost the ability to interact with the 'anticlockwise' side of Vma16. The other paralogue (which became Vma11) lost the ability to interact with the 'clockwise' side of Vma16 and/or the anticlockwise side of Vma3.

A general theory, sometimes called constructive neutral evolution, to explain how neutral processes might drive a system towards complexity is more than a decade old[8-10]. But the study by Thornton and colleagues[1] may provide the most compelling experimental evidence to date. Of course, one can never prove that some subtle, unidentified selective advantage was not involved in the evolution of the V-ATPase protein ring, but neutrality would seem the most justifiable default hypothesis. Thus, a neutral theory of molecular evolution, normally invoked for nucleotide substitutions, may also apply to certain higher-order structures such as multi-subunit protein rings. How general such neutral mutational drives to complexity might be is one of evolutionary theory's deeper unanswered questions. ■

**W. Ford Doolittle** *is in the Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada.*
*e-mail: ford@dal.ca*

1. Finnigan, G. C., Hanson-Smith, V., Stevens, T. H. & Thornton, J. W. *Nature* **481,** 360–364 (2012).
2. Lynch, M. *Proc. Natl Acad. Sci. USA* **104,** 8597–8604 (2007).
3. Archibald, J. M., Logsdon, J. M. Jr & Doolittle, W. F. *Curr. Biol.* **9,** 1053–1056 (1999).
4. Ruano-Rubio, V. & Fares, M. A. *Mol. Biol. Evol.* **24,** 1384–1396 (2007).
5. Fares, M. A. & Wolfe, K. H. *Mol. Biol. Evol.* **20,** 1588–1597 (2003).
6. Ohno, S. *Evolution by Gene Duplication* (Springer, 1970).
7. Wang, Y., Cipriano, D. J. & Forgac, M. *J. Biol. Chem.* **282,** 34058–34065 (2007).
8. Stoltzfus, A. *J. Mol. Evol.* **49,** 169–181 (1999).
9. Force, A. *et al. Genetics* **151,** 1531–1545 (1999).
10. Gray, M. W., Lukeš, J., Archibald, J. M., Keeling, P. J. & Doolittle, W. F. *Science* **330,** 920–921 (2011).

GALAXY FORMATION

# Distant dwarfs

**Astronomers value the smallest galaxies for the clues they can provide about how galaxies form. But only those nearest to us are easily detectable. A neat technique has allowed one such object to be found at a large distance.** SEE LETTER P.341

ROBERT W. SCHMIDT

Some objects do not shine brightly enough to be seen easily. Examples of this are the small and extremely faint galaxies called dwarf or satellite galaxies. Only when they are near to our Galaxy might we be lucky enough to detect the light of their stars. Satellite galaxies are of great interest because simulations of the process of galaxy formation predict that there are many more satellites in the Local Group, which includes the Milky Way and other galaxies in its vicinity, than have yet been found[1-3]. This begs the question of whether there is something wrong with our understanding of how galaxies form. On page 341 of this issue, Vegetti *et al.*[4] report the discovery of a satellite in the vicinity of a galaxy located at a cosmological distance from Earth that allows them to test the simulation predictions for galaxies beyond those of the Local Group.

The object studied by Vegetti *et al.* is a distant elliptical galaxy that acts as a gravitational lens: light from a background source is deflected from its otherwise straight path by a massive foreground object, the lens. When compared with camera lenses, gravitational lenses have rather dismal optical properties because they can lead to large distortions and multiple images of the background object. But from an astronomer's point of view this is a good thing: the mass of the gravitational lens can be measured by reconstructing the background object from the distorted images.

## 50 Years Ago

At the International Geological Congress in Copenhagen in 1960, and at the Russian Trade Exhibition in London last year, one of the outstanding scientific exhibits was a hand-coloured tectonic map of Europe compiled by Russian draughtsmen from copy submitted by the various national geological surveys. The map, on a scale of 1:2,500,000, has been sponsored by the Sub-Commission on the Tectonic Map of the World of the International Geological Congress … Western geologists who wish to make sure of receiving a copy should place an order, in advance of publication, with their national agents for Russian books.
**From *Nature* 20 January 1962**

## 100 Years Ago

The possibility of the discovery of a remedy for cancer has been advanced a stage by the preparation by Prof. Wassermann, of Berlin, of a substance which possesses a curative action experimentally on cancer of mice. Prof. Wassermann reasoned that since the cancer-cells are growing rapidly, their oxygen requirements would be different from, and greater than, those of the cells of the body generally. He sought for some substance which might interfere with the oxygen supply to the cancer-cells, and finally adopted selenium as a means to do this. The next problem was to convey selenium to the cancer-cells by means of the blood stream, and … a compound of selenium with an anilin dye eosin was found to fulfil this condition … After two or three injections of the substance into a mouse the subject of cancerous tumours, the tumours are found to have softened, and after six to eight doses they become cystic, diminish in size, and finally disappear, and no recurrence takes place. The eosin–selenium compound is, however, poisonous, and a certain number of mice succumb under the treatment.
**From *Nature* 18 January 1912**

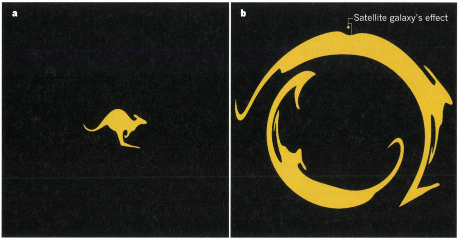


**Figure 1 | Seeking satellites.** Vegetti *et al.*[4] have discovered a satellite galaxy around a larger galaxy, located at a cosmological distance, by examining how both galaxies act as gravitational lenses and distort light from a background source as the light travels to Earth. **a**, Were there no lenses between an observer on Earth and a source — here illustrated as a kangaroo[10] — the observer would, if instrumentation allowed, see one image of the source. **b**, If there is a massive lens between the observer and the source, the source is magnified and imaged more than once. In this case, as in the authors' study, the source is deformed into a ring that consists mainly of two distorted images of the source. The two images are caused by the larger galaxy, whereas the dent and blob features (labelled) are caused by the satellite galaxy, the mass of which has been exaggerated to demonstrate the effect more clearly.

In the present study, the gravitational lens system, called JVAS B1938 + 666, consists of a very distant galaxy that bends light from an even more distant background galaxy. Light from the background galaxy is deformed into a ring around the lens galaxy — a prime example of what is known as an Einstein ring. In this case, the ring is formed mainly by two lensed images of the background galaxy[5].

Vegetti and her team[4] obtained a near-infrared image of JVAS B1938 + 666 using the 10-metre Keck telescope in Hawaii. They used an optics system that corrects for the blurring effect of Earth's atmosphere to improve the image quality. With such data, they could neatly determine the mass distribution of the lens galaxy, as well as the shape and brightness of the background galaxy.

And here comes the connection to satellite galaxies. The sophisticated numerical technique[6] used by the authors allowed them not only to derive a model of the lens galaxy's mass, but also to map any excess lens mass that could not be accounted for by the galaxy. They found an excess mass near the Einstein ring that they attributed to the presence of a satellite galaxy (Fig. 1). This method has been dubbed gravitational imaging[6]. Vegetti *et al.* also used an analytical model to test the detected excess mass and found that a satellite galaxy is indeed required to explain the data.

This satellite galaxy is exciting because it was detected in the excess-mass map despite its low mass. Assuming that the object is in the neighbourhood of the main lens galaxy, it has a mass of some 113 million solar masses within a radius of about 600 parsecs (2,000 light years) — values that put it firmly in the realm of satellite galaxies. The mass is tenfold lower than those of two other satellite galaxies that have been detected[7,8] recently using gravitational lenses (with the possible exception of another low-mass satellite[8], although the detection of this object is not yet confirmed).

Vegetti and colleagues[4] went on to determine the mass function of satellite galaxies — that is, the expected number of satellites for a given mass — at distances beyond the local Universe by combining data from the new object and another satellite galaxy[7]. They found that the resulting mass function is consistent with current galaxy-formation theory. But owing to the small number of known objects, uncertainties in the deduced function are quite large.

A natural question to ask is whether the satellite galaxy can be observed directly rather than by its gravitational effect on the shape of a background object. With current instrumentation, the answer is no. The object is simply too distant to be imaged directly. But the message here[4] is that it is possible to spot these elusive objects around distant lens galaxies without knowing where to look for them.

The satellite should be considered in the context of about 50 satellite galaxies that are known to exist in the Local Group and that have been found by imaging their starlight[3,9]. The number of these nearby satellites has greatly increased over the past decade, and often only a hundred (or fewer) stars are detected in them. From measurements of the velocities of their member stars, the satellites have been shown to be dominated by an unseen mass component called dark matter. Dark matter is the reason that satellites such as the newly discovered one are massive enough to be found using gravitational lenses.

Despite this increase in the observed number of satellites in the Local Group, simulations indicate that there are many more satellites dominated by dark matter than have been seen. But a fair fraction of the local satellites will probably be found by sensitive searches[3], and theorists have plenty of ideas about why satellites may be lacking in stars[1,2]. If more satellites are discovered at cosmological distances using gravitational lenses, they will have the potential to independently improve our knowledge of the satellite mass function[4,7,8].

Vegetti and colleagues show convincingly that gravitational lenses allow low-mass satellite galaxies to be discovered regardless of how luminous the satellites are. As such, they offer the means to refine measurements of the mass function of faraway satellites. ■

---

**Robert W. Schmidt** *is at the Zentrum für Astronomie der Universität Heidelberg, Astronomisches Rechen-Institut, 69120 Heidelberg, Germany.*
*e-mail: rschmidt@ari.uni-heidelberg.de*

1. Bullock, J. S. Preprint at http://arxiv.org/abs/1009.4505 (2010).
2. Kravtsov, A. *Adv. Astron.* **2010,** 281913 (2010).
3. Willman, B. *Adv. Astron.* **2010,** 285454 (2010).
4. Vegetti, S. *et al. Nature* **481,** 341–343 (2012).
5. King, L. J. *et al. Mon. Not. R. Astron. Soc.* **295,** L41–L44 (1998).
6. Koopmans, L. V. E. *Mon. Not. R. Astron. Soc.* **363,** 1136–1144 (2005).
7. Vegetti, S. *et al. Mon. Not. R. Astron. Soc.* **408,** 1969–1981 (2010).
8. Fadely, R. & Keeton, C. R. *Mon. Not. R. Astron. Soc.* **419,** 936–951 (2012).
9. Richardson, J. C. *et al. Astrophys. J.* **732,** 76 (2011).
10. Wambsganss, J. *Nature* **386,** 27–28 (1997).

---

CELL CYCLE

# A division duet

**The orchestration of cell division requires a programme of events choreographed by protein modification. A study shows that the relative activity of a phosphatase enzyme towards its substrates imposes order during the final act of division.**

CURT WITTENBERG

Protein phosphorylation plays a central part in the intricate choreography of cell-cycle events[1]. A minimalistic model that holds that simple oscillations in the activity of a kinase enzyme called CDK — which phosphorylates cell-cycle proteins — is sufficient to promote the orderly progression of the cycle has been tested and confirmed[2]. However, protein phosphorylation, like other reversible cellular processes, depends not only on enzymes that catalyse the addition of phosphate groups, but also on those that catalyse dephosphorylation — the protein phosphatases. In a paper published in *Cell*, Bouchoux and Uhlmann[3] show that the intrinsic ability of a protein phosphatase to dephosphorylate specific substrates is crucial for establishing the proper order of events during the final stages of mitotic cell division*.

The cell cycle proceeds from the G1 gap phase through the S phase, in which chromosomes are duplicated, and then on to the G2 gap phase and M phase (mitosis). At this point, the duplicated chromosomes are separated, leading to the creation of two daughter cells back at G1 phase. The protein phosphatase Cdc14 is required for the completion of chromosome segregation and subsequent events that lead to the resetting of the 'cell-cycle clock' such that the G1-phase daughter cells can initiate a new cycle[4] (Fig. 1). Those events are prevented before M phase because Cdc14 is sequestered in a region of the nucleus called the nucleolus, and so does not have access to its protein targets. However, it has remained unclear how, after its release, Cdc14 triggers the events that it regulates in a specific order. Indeed, elongation of the mitotic spindle occurs subsequent to spindle attachment to chromosomes, but before cell separation during cytokinesis, the final stage of mitosis.

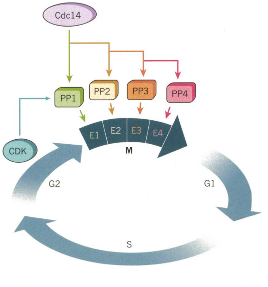


**Figure 1 | Imposing order through protein modification.** The cell cycle proceeds from the G1 phase to the S, G2 and, finally, M phases. At the end of M phase, the cell divides into two daughter cells, which enter the G1 phase. Several events (hypothetical events E1–E4) mediate the passage from M phase into G1 phase, and, for them to occur in an orderly manner, proteins phosphorylated by CDK (hypothetical phosphoproteins PP1–PP4) must be dephosphorylated by Cdc14 in the right order (from PP1 to PP4). Bouchoux and Uhlmann[3] show that the differential activity of Cdc14 towards CDK-phosphorylated proteins determines their orderly dephosphorylation, and so the order of events that depend on them.

Potential resolutions of that conundrum were set out in early models that explain the order of events during the cell cycle. In the domino model, order is a consequence of the dependency of later events on completion of those occurring earlier; in the clock model, order is imposed by an intrinsic timer[5]. Bouchoux and Uhlmann's study distinguishes between these two models, at least in the context of cell-cycle events that depend on Cdc14.

The authors engineered yeast strains in which the activities of Cdc14 and CDK could be modulated. Using careful *in vivo* analysis, they found that Cdc14 dephosphorylates CDK substrates in a sequence consistent with the order of cell-cycle events. That is, dephosphorylation events that promote spindle elongation occur after those that promote spindle attachment to chromosomes, but before those that promote spindle disassembly.

Cdc14 activity rises upon release from the nucleolus, allowing it to compete with CDK activity[6]. Bouchoux and Uhlmann[3] performed careful analysis using purified proteins *in vitro* and show that the timing of dephosphorylation *in vivo* is a consequence of Cdc14's differential activity towards specific substrates. That is, 'early' Cdc14 substrates — those dephosphorylated *in vivo* during the initial Cdc14-triggered events — are more efficiently dephosphorylated by this enzyme *in vitro* than are 'late' substrates (Fig. 1). Furthermore, by establishing a competition between CDK and Cdc14 *in vitro* — a situation that exists during anaphase, the penultimate stage of mitosis, when both enzymes are highly active — they showed that dephosphorylation of early substrates is favoured. These findings establish that differences in the intrinsic activity of Cdc14 for specific substrates is sufficient to explain the order of dephosphorylation events observed *in vivo*; they also show that Cdc14 is an important component of a cell-cycle clock.

Could mechanisms besides the intrinsic activity of Cdc14 towards its substrates determine the order of protein dephosphorylation? Bouchoux and

---

*This article was published online on 11 January 2012.

Uhlmann go to some lengths to exclude the involvement of Cdc14 localization within the nucleus, the dependence of dephosphorylation on specific events associated with mitotic exit, and the specificity that results from the activation of CDK by its different activating partners (cyclins).

Uhlmann and colleagues[7] have previously suggested that a CDK-based oscillator acting as a counterbalance to protein phosphatases might be responsible for order not just during mitotic exit, but throughout cell division. However, there is substantial biochemical and genetic evidence for interdependent relationships that impose order throughout the cell cycle[6]. For example, a recent study[8] demonstrated that the sequential action of specific cyclin–CDK complexes is required to eliminate Sic1 — a potent CDK inhibitor that restricts passage from G1 into S phase. The researchers[8] report that the order of action of those cyclin–CDK complexes is due not to differences in their intrinsic activity, but to differences in their specificity for particular phosphorylation sites on Sic1. That is, phosphorylation of a set of sites by one cyclin–CDK complex depends on prior phosphorylation of Sic1 by a different cyclin–CDK complex.

It seems, therefore, that multiple mechanisms are involved in imposing order on cell-cycle events. Nevertheless, the finding[2] that a single cyclin–CDK complex can orchestrate an effective cell-division cycle suggests that, at least in some systems and under certain physiological conditions, the specificity of different cyclin–CDK complexes is dispensable. It remains to be established whether phosphatase-based mechanisms similar to those described by Bouchoux and Uhlmann[3] are essential for ordering cell-cycle events beyond mitotic exit, and whether they are even sufficient to replace cyclin–CDK specificity under particular conditions. It is largely because the other phosphatases that provide the counterpoint to CDK are poorly understood that the choreography of this intricate dance remains incompletely described. ■

**Curt Wittenberg** *is in the Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92024, USA.*
*e-mail: curtw@scripps.edu*

1. Morgan, D. O. *Annu. Rev. Cell Dev. Biol.* **13,** 261–291 (1997).
2. Coudreuse, D. & Nurse, P. *Nature* **468,** 1074–1079 (2010).
3. Bouchoux, C. & Uhlmann, F. *Cell* **147,** 803–814 (2011).
4. Queralt, E. & Uhlmann, F. *Curr. Opin. Cell Biol.* **20,** 661–668 (2008).
5. Murray, A. W. & Kirschner, M. W. *Science* **246,** 614–621 (1989).
6. Sullivan, M. & Morgan, D. O. *Nature Rev. Mol. Cell Biol.* **8,** 894–903 (2007).
7. Uhlmann, F., Bouchoux, C. & López-Avilés, S. *Phil. Trans. R. Soc. Lond. B* **366,** 3572–3583 (2011).
8. Kõivomägi, M. *et al. Nature* **480,** 128–131 (2011).

---

ORGANIC CHEMISTRY

# Stabilizers cause instability

**Any chemist will tell you that amide compounds are renowned for their stability, yet a family of amides has been discovered that is surprisingly unstable. The cause of this phenomenon is even more surprising.**

**JONATHAN CLAYDEN**

Some molecules are stable under all kinds of conditions; others are so unstable that they'll react with almost anything. Amides, the class of compounds that includes proteins and nylon, are usually stable molecules. Strong acid or alkali is needed to break them down — nylon fabrics are highly resistant to chemical damage, and dietary proteins can be digested only by the acidity of the stomach coupled with a battery of enzymes. Writing in *Angewandte Chemie*, Hutchby *et al.*[1] report a group of amides that turns this typical character on its head. These compounds share structural characteristics with the least reactive amides of all, yet a remarkable mechanistic quirk means they exhibit reactivity that is usually associated with unstable reagents.

In most cases, the reactivity of a molecule can be predicted with a fair degree of certainty by considering the combined effects of bond strength, polarity and steric hindrance (the physical volume occupied by parts of a molecule). For example, a carbon–iodine bond is more reactive than a carbon–fluorine bond because it is much weaker, whereas an oxygen–hydrogen bond is more reactive than a carbon–hydrogen bond because it is more polarized. These effects mean that the most interesting properties of organic molecules tend to be concentrated in reactivity hotspots known as functional groups, each of which contains one or more atoms that either attract or repel electrons more than does carbon or hydrogen. The concept of the functional group, and the fact that functional groups have a typical reactivity profile, is central to the ability of chemists to assess, predict and design the properties of molecules.

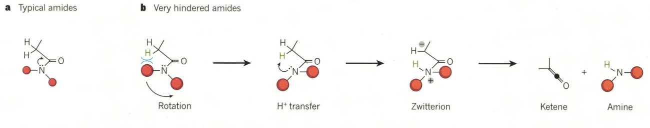


**Figure 1 | Amide reactivity.** **a,** Most amides are unreactive because the nitrogen atom in the compounds donates a pair of electrons (shown as dots) to the carbonyl (C=O) group, and because the functional groups (red spheres) on the nitrogen atom may block the approach of attacking reagents. The curly arrow indicates electron donation. **b,** Hutchby *et al.*[1] report that very hindered amides, which have outsize groups on the nitrogen atom, buck the trend, because the groups are repelled (blue curved lines) by other groups in the molecule. This repulsion facilitates rotation of the bond between the nitrogen and the carbon of the carbonyl group, and so of the groups on the nitrogen atom, and prevents donation of the nitrogen's electron pair to the carbonyl group. Instead, a hydrogen atom (green) on the carbon adjacent to the carbonyl group is transferred as a $H^+$ ion to the nitrogen, forming a zwitterion (a molecule that contains both a positive and a negative charge), which falls apart to produce a ketene and an amine. The black dot in the ketene represents a carbon atom.

One particular family of functional groups — those that contain a carbonyl group, C=O — is of particular importance because its reactions are central to most common methods of constructing carbon–carbon bonds. The carbonyl group is electrophilic (it reacts with electron-rich species such as alcohols and amines), but just how electrophilic it is depends enormously on what else is attached to the group's carbon atom. Attach an electron-withdrawing chlorine atom, and a highly reactive acid chloride results; attach two chlorine atoms and you get phosgene, an even more reactive (and toxic) gas. Appending an electron-rich nitrogen atom, on the other hand, satisfies the carbonyl group's desire for electrons because the nitrogen shares its own electron pair with the group (Fig. 1a), a phenomenon known as delocalization. The resulting compound is an unreactive amide.

Amides are unreactive not just because the nitrogen atom satisfies the electron deficiency of the carbonyl group, but also because the nitrogen atoms of amides can carry two substituents, both of which can get in the way of incoming reagents (the phenomenon of steric hindrance; Fig. 1a). Hindered amides, in which the nitrogen atom carries two big groups, are usually among the least reactive carbonyl compounds known. This makes them valuable in organic chemistry — with the reactivity of the carbonyl group shut down, reactions can occur elsewhere without interference from competing side reactions at the carbonyl group. Hindered amides are used in this way in the preparation of reactive, synthetically useful organometallic compounds[2], for example, and can be even less reactive towards electron-rich reagents than is the proverbially stable benzene ring[3,4].

The group of compounds now reported by Hutchby *et al.*[1] overturns these expectations. The authors prepared highly encumbered amides in which the nitrogen atom carries two very large groups, and observed that, far from being the least reactive of their class, the compounds were more reactive than their less encumbered relatives. A possible explanation for this is an effect that takes over from steric hindrance when the groups carried by an amide's nitrogen atom are abnormally large: the groups effectively switch off the carbonyl-deactivating effect of the nitrogen's pair of electrons (Fig. 1b).

Such 'non-delocalized' amides, in which the nitrogen's electrons are forced into a position from which they cannot interact with the carbonyl group, have been made before[5,6], and are indeed more reactive. But Hutchby *et al.* propose[1] that the 'off switch' is more subtle in their case. They found that these amides have structural characteristics very similar to unhindered amides, but that the repulsion between the nitrogen's large groups and other parts of the molecule makes it much more favourable for the groups to move away from their neighbouring atoms. This allows the bond between the carbonyl group and the nitrogen atom to rotate (Fig. 1b), in turn allowing the nitrogen atom to take back its previously delocalized, deactivating electrons. Those electrons therefore become available to grab hold of a hydrogen ion ($H^+$) from the other side of the carbonyl group, setting in train a series of events — again accelerated by the size of the groups on the nitrogen — in which the molecule is forced to fragment into a compound called a ketene, which is much more reactive than an amide. In the authors' experiments, the ketene subsequently reacted with the surrounding solvent (methanol), forming ester compounds as the final products.

Not all highly encumbered amides will behave as Hutchby and colleagues' compounds do, only those that have a methylene group ($CH_2$) on the other side of the carbonyl group from the nitrogen atom. And for the process to work really well, an electron-withdrawing functional group must be attached to the methylene group. Nevertheless, it is remarkable that the least reactive of carbonyl compounds can so easily be bumped up the reactivity scale, and even more so that the very change responsible — increased molecular bulk — usually decreases reactivity. This discovery exposes a chink in the armour of amides that could be used to circumvent many current and future challenges in synthetic chemistry. For example, amides could be used as a Trojan horse to incorporate hidden reactivity into a molecule, reactivity that is later revealed using Hutchby and colleagues' chemistry when other modifications to the molecule have been made. ■

**Jonathan Clayden** *is in the School of Chemistry, University of Manchester, Manchester M13 9PL, UK.*
*e-mail: clayden@man.ac.uk*

1. Hutchby, M. *et al. Angew. Chem. Int. Edn* **51,** 548–551 (2012).
2. Snieckus, V. *Chem. Rev.* **90,** 879–933 (1990).
3. Clayden, J., Foricher, Y. J. Y. & Lam, H. K. *Chem. Commun.* 2138–2139 (2002).
4. Clayden, J., Foricher, Y. J. Y. & Lam, H. K. *Eur. J. Org. Chem.* 3558–3565 (2002).
5. Clayden, J. & Moran, W. J. *Angew. Chem. Int. Edn* **45,** 7118–7120 (2006).
6. Szostak, M. & Aubé, J. *Org. Biomol. Chem.* **9,** 27–35 (2011).

---

CANCER

# Reprogramming clinical outcome

**On binding to its target hormone, the oestrogen-receptor protein modulates the expression of many genes. Changes in the receptor's interaction with DNA have now been linked to clinical outcome in patients with breast cancer.** **SEE LETTER P.389**

**GENEVIÈVE DEBLOIS & VINCENT GIGUÈRE**

In about two-thirds of breast cancers, tumour growth depends on oestrogen. This hormone binds to and activates its nuclear receptor, which then binds to specific sites on the genome and modifies gene expression. Hormonal therapies that inhibit oestrogen synthesis or that block the oestrogen receptor with drugs such as tamoxifen can restrain the proliferation of cancer cells in oestrogen-dependent breast tumours. Consequently, the presence of this receptor and the expression of its target genes in primary tumours are generally predictors of a good prognosis, as the tumour will be amenable to these treatments. But despite the sustained presence of the receptor, these tumours can eventually develop resistance to hormonal therapies and spread to distant sites, leading to a poor outcome. On page 389 of this issue, Ross-Innes *et al.* (Carroll and colleagues)[1] present evidence that reprogramming of the interaction of the oestrogen receptor with DNA in breast tumours significantly contributes to this change in clinical outcome.

Functional genomics — genome-wide studies of dynamic interactions between genes and proteins — has greatly enhanced our understanding of the action of the oestrogen receptor (ERα) in breast cancer[2]. Such analyses have identified a total of some 80,000 ERα-binding sites in the genomes of breast-cancer cells. However, there is only a relatively modest overlap between the sets of binding sites described by different studies. This indicates that, with the exception of a small set of core binding sites, plasticity is an attribute of ERα recruitment to chromatin (the complex of DNA and proteins). So, could clinicians use the ERα-binding profile of a breast tumour to predict the cancer's fate?

To address this question, Ross-Innes *et al.*[1] mapped the ERα binding sites in the genomes of tumour samples by sequencing DNA segments to which ERα was bound. The results clearly establish that different ERα-binding profiles in primary breast tumours can be

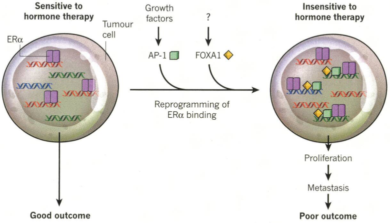

Sensitive to hormone therapy
Tumour cell
ERα
Growth factors
?
AP-1
FOXA1
Reprogramming of ERα binding
Insensitive to hormone therapy
Proliferation
Metastasis
Good outcome
Poor outcome

**Figure 1 | Reprogramming of oestrogen-receptor activity in breast cancer.** Tumour cells that express the oestrogen receptor ERα are sensitive to hormone-based therapy, leading to a good therapeutic outcome. Carroll and colleagues[1] show that ERα binding to DNA is dynamically regulated, and that different ERα-binding profiles are associated with different clinical outcomes. They propose that the transcription factor FOXA1 can mediate the reprogramming of ERα activity by facilitating the recruitment of ERα to different locations on DNA, leading to the development of resistance to hormonal therapy and a bad therapeutic outcome. The biological cues acting on FOXA1 to induce ERα reprogramming are unknown. Another transcription factor, AP-1, might also reprogram ERα activity in breast tumours. Signalling by growth factors has been linked to resistance to hormonal therapy[7] and is known to activate AP-1. Ross-Innes *et al.*[1] observe a comparable presence of sites for AP-1 and FOXA1 binding to DNA in ERα-bound DNA segments identified in tumours with poor outcome.

linked to distinct clinical outcomes. Thus, profiling of ERα-binding sites might contribute to a better molecular classification of breast tumours.

In the current study, the authors pre-classified breast tumours from 15 patients into two groups — good and poor prognosis — depending on the presence of certain proteins that are used as standard markers for responsiveness to hormonal therapy. On its own, this approach can be deceptive, because some tumours that express markers predicting high responsiveness show only a modest response to tamoxifen and therefore have a poor outcome. So, in their pre-classification, Ross-Innes *et al.*[1] also took into consideration information from long-term follow-up of the patients. However, to establish precise binding profiles that, by themselves, could be used to predict outcome without prior knowledge of other indicators, many more tumours from larger groups of diverse patients would have to be profiled.

What mechanism underlies the differences among breast tumours in ERα recruitment to genomic sites? Previous functional genomic studies[3,4] have shown that the interaction of ERα with chromatin largely depends on the presence of the protein FOXA1. This belongs to a class of transcription factors that facilitate protein binding to chromatin[5]. Recently, Carroll and colleagues[6] argued that FOXA1 is essential for ERα binding to chromatin, and for the receptor's subsequent transcriptional activity, not only in the presence of oestrogen but also in cells exposed to tamoxifen. Whether FOXA1 plays a part in the response of ERα-positive breast tumours to hormonal therapies in patients remained unclear, however.

In the absence of co-profiles of FOXA1 and ERα binding in the same breast-tumour samples, Ross-Innes *et al.*[1] addressed this question by searching within ERα-binding regions for known DNA sequence motifs recognized by different transcription factors. Unexpectedly, they found that the FOXA1-binding motif occurs on ERα-bound DNA segments obtained from tumours associated with a poor outcome but not on segments from tumours associated with a good outcome.

To validate this observation, the authors used cell-based models of breast cancer and again found preferential enrichment of the FOXA1-binding motif in tamoxifen-resistant cells and in cells that had been treated with a cocktail of agents promoting growth and proliferation. They conclude that in the tumours associated with poor outcome, FOXA1 could be triggering a rapid and dynamic reprogramming of ERα recruitment to chromatin at different sites from those in tumours associated with a good outcome (Fig. 1).

Whether FOXA1-dependent programming of ERα binding to DNA occurs during the course of the disease or is intrinsic to the primary tumour is yet to be determined. In either case, it would be of interest to identify the physiological cues and mechanisms dictating FOXA1 recruitment to chromatin. Identifying FOXA1-binding sites in breast-tumour samples would be an essential first step in assessing these issues.

On the basis of their previous work[6], Carroll and colleagues logically favour the hypothesis that FOXA1 has a key role in mediating reprogramming of ERα activity. It is likely, however, that additional mechanisms are involved, including the participation of other proteins. In particular, AP-1, a transcription factor that stimulates cell growth and proliferation, might also contribute to ERα reprogramming (Fig. 1). Indeed, the authors[1] find an enrichment of the binding motif for AP-1 in ERα-bound DNA segments from tumours associated with poor outcome. This observation nicely parallels previous results[7] showing enrichment of AP-1 motifs in ERα-bound segments upon treatment of breast cancer cells with epidermal growth factor, a condition that mimics the increase in growth-factor signalling in hormone-resistant tumours. The involvement of AP-1 in this process is also supported by its function in facilitating chromatin availability for the binding of the glucocorticoid receptor — another nuclear hormone receptor — to chromatin[8].

Over the past few years, functional genomics has enhanced our knowledge of hormone action. Ross-Innes and co-workers' technical prowess[1] should motivate similar studies of other nuclear receptors and their interacting protein partners involved in the initiation and progression of all subtypes of breast cancer[9–12]. Profiling of transcription factors in cancer-cell lines has proved extremely useful in deciphering the mechanisms underlying hormone-driven gene expression. But using primary tumours and metastatic tissues, as Ross-Innes *et al.* have done, will be required to reach relevant conclusions about the roles of each of these factors in cancer. ■

**Geneviève Deblois** *and* **Vincent Giguère** *are at the Goodman Cancer Research Centre and Departments of Biochemistry, Medicine and Oncology, McGill University, Montreal, Quebec H3A 1A3, Canada.*
*e-mail: vincent.giguere@mcgill.ca*

1. Ross-Innes, C. S. *et al. Nature* **481,** 389–393 (2012).
2. Lupien, M. & Brown, M. *Endocr.-Relat. Cancer* **16,** 381–389 (2009).
3. Carroll, J. S. *et al. Cell* **122,** 33–43 (2005).
4. Laganière, J. *et al. Proc. Natl Acad. Sci. USA* **102,** 11651–11656 (2005).
5. Zaret, K. S. & Carroll, J. S. *Genes Dev.* **25,** 2227–2241 (2011).
6. Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. *Nature Genet.* **43,** 27–33 (2011).
7. Lupien, M. *et al. Genes Dev.* **24,** 2219–2227 (2010).
8. Biddie, S. C. *et al. Mol. Cell* **43,** 145–155 (2011).
9. Robinson, J. L. *et al. EMBO J.* **30,** 3019–3027 (2011).
10. Ni, M. *et al. Cancer Cell* **20,** 119–131 (2011).
11. Deblois, G. *et al. Cancer Res.* **69,** 6149–6157 (2009).
12. Xu, J., Wu, R. C. & O'Malley, B. W. *Nature Rev. Cancer* **9,** 615–630 (2009).

nature INSIGHT

# FRONTIERS IN BIOLOGY

REDISCOVER *your* nature

browse.
discover.
Save.
READ.
3G
9:41 AM
Back
nature.com
NEWS
Domino synthesis mimics nature's chemistry
Nature News
NEWS
A kinder, gentler defibrillator
Nature News
NEWS
The youngest dinosaur fragment yet?
Nature News
NEWS
Recent articles
Saved
Subscriptions
Sea

## ABSTRACTS AND ARTICLES

Now you can browse and search abstracts from *Nature* and selected other nature.com journals on the go with the nature.com Reader app, for iPhone® or iPad®. Save articles to your nature.com account for later reading on your PC or Mac. Access to abstracts and open access content is free through the app; other full-text articles are available through relevant personal or mobile subscriptions. **Search for 'nature.com' on the App Store.**

Apple, the Apple logo, iPhone, iPad and iPod touch are trademarks of Apple Inc., registered in the U.S. and other countries.

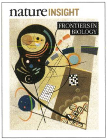
Available on the App Store

**nature.com/mobileapps**

# nature INSIGHT

## FRONTIERS IN BIOLOGY

19 January 2012 / Vol 481 / Issue No 7381



**Cover illustration**
by Nik Spencer

**Editor, *Nature***
Philip Campbell
**Publishing**
Nick Campbell
**Insights Editor**
Ursula Weiss
**Production Editor**
Nicola Bailey
**Art Editor**
Nik Spencer
**Sponsorship**
Gerard Preston
**Production**
Emilia Orviss
**Marketing**
Elena Woodstock
Hannah Phipps
**Editorial Assistant**
Hazel Mayhew

**The Macmillan Building**
**4 Crinan Street**
**London N1 9XW, UK**
**Tel: +44 (0) 20 7833 4000**
**e: nature@nature.com**

nature publishing group

The Nature Insight 'Frontiers in Biology' aims to cover timely and important developments in biology, ranging from the subcellular to the organismal level, and including molecular mechanisms and biomedicine. In this Insight the reviews discuss the role of innate immune signalling in tissue homeostasis and the response to infection, how understanding the DNA damage response has guided the development of inhibitors and helped to establish new principles for treating cancer, the potential therapeutic promise of patient-derived pluripotent stem cells, the forces that govern tumour evolution and the impact of mouse genetics on the study of bone physiology.

Richard Flavell and colleagues provide a timely overview of the activation and function of different inflammasomes. These are protein complexes that sense microbes and signs of tissue damage by responding to exogenous and endogenous signals. They can be beneficial and detrimental in a diverse set of inflammatory and metabolic diseases.

Christopher Lord and Alan Ashworth discuss how understanding the different response pathways to DNA damage has led to a multipronged approach to cancer therapy. Inherent mutations of the cancer cell and repair-enzyme inhibitors are used to synergistically target multiple repair pathways to enhance the cell killing effect.

Daisy Robinton and George Daley highlight the recent progress in the derivation of induced pluripotent stem cells, outline the functional assessments of pluripotency and discuss whether induced pluripotent stem cells equal embryonic stem cells, and thus may be a suitable alternative for research and therapy.

Mel Greaves and Carlo Maley consider our current knowledge of clonal evolution in cancer. Tumours and metastases evolve through the successive accumulation of genetic and epigenetic alterations and clonal selection within the adaptive environment. This leads to phenotypic and functional heterogeneity of cancer cells within a tumour. Although therapy can eliminate most cancer cells, it can drive the selection of resistant clones.

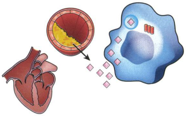
Finally, Gérard Karsenty and Mathieu Ferron discuss how recent evidence has changed our view of the skeleton and revealed an intimate connection between bone and the regulation of metabolism.

**Alex Eccleston, Angela Eggleston, Marie-Thérèse Heemels, Barbara Marte & Ursula Weiss**
*Senior Editors*

## CONTENTS

### REVIEWS

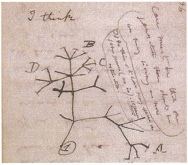
278 **Inflammasomes in health and disease**
*Till Strowig, Jorge Henao-Mejia, Eran Elinav & Richard Flavell*



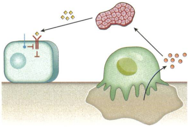
287 **The DNA damage response and cancer therapy**
*Christopher J. Lord & Alan Ashworth*

295 **The promise of induced pluripotent stem cells in research and therapy**
*Daisy A. Robinton & George Q. Daley*

306 **Clonal evolution in cancer**
*Mel Greaves & Carlo C. Maley*


I think
B
D
C
1
A

314 **The contribution of bone to whole-organism physiology**
*Gérard Karsenty & Mathieu Ferron*



doi:10.1038/nature10759

# Inflammasomes in health and disease

Till Strowig[1]*, Jorge Henao-Mejia[1]*, Eran Elinav[1]* & Richard Flavell[1,2]

**Inflammasomes are a group of protein complexes built around several proteins, including NLRP3, NLRC4, AIM2 and NLRP6. Recognition of a diverse range of microbial, stress and damage signals by inflammasomes results in direct activation of caspase-1, which subsequently induces secretion of potent pro-inflammatory cytokines and a form of cell death called pyroptosis. Inflammasome-mediated processes are important during microbial infections and also in regulating both metabolic processes and mucosal immune responses. We review the functions of the different inflammasome complexes and discuss how aberrations in them are implicated in the pathogenesis of human diseases.**

Inflammation is an acute response to infection and tissue damage to limit harm to the body[1]. However, dysregulated and chronic inflammation may result in secondary damage and immune pathology to the host. Inflammation is initiated on the sensing of signs of acute damage or disturbances of the steady state. Several recognition systems have evolved together to distinguish between homeostasis and threats to the host. Some of these receptors recognize distinct conserved pathogen-associated molecular patterns (PAMPs) that are predominantly found in microbes and hence allow the exquisite sensing of pathogens in tissues that are normally devoid of these structures[2]. Multiple additional receptors recognize host-derived signals, called damage-associated molecular patterns (DAMPs), that are released as a result of perturbations of tissue homeostasis caused by microbial or non-microbial insults, allowing a general sensing of stressed tissue[3].

Inflammasomes are a group of protein complexes that recognize a diverse set of inflammation-inducing stimuli that include PAMPs and DAMPs and that control the production of important pro-inflammatory cytokines such as interleukin-1β (IL-1β) and IL-18 (refs 4, 5). Furthermore, they have been found to regulate other important aspects of inflammation and tissue repair such as pyroptosis, a form of cell death. Recent studies have characterized distinct molecular activation mechanisms for several sensor proteins and have identified a multitude of ligands of both endogenous and exogenous origins. Moreover, the diverse functions of these complexes in antimicrobial responses, as well as in multifaceted diseases such as metabolic syndrome and inflammatory bowel disease (IBD), have started to be revealed. Importantly, mutations in components of inflammasome complexes have been associated with a propensity for the development of several immune-mediated diseases in humans. We review several of these conditions, discuss the different models that have been proposed for inflammasome involvement in normal and aberrant immune response, and highlight the challenges and future directions for this field.

## Inflammasomes are intracellular innate immune sensors

### Inflammasomes control the activity of caspase-1

IL-1β is one of the quintessential pro-inflammatory cytokines that broadly affects inflammatory processes[6]. Tight control of its production is therefore required at the transcriptional and post-translational levels. IL-1β is synthesized as a pro-protein without a typical signal sequence that would allow its secretion, and instead its activation and cellular release are controlled by the cysteine protease caspase-1 (ref. 7). Similarly, caspase-1 is responsible for the processing and secretion of IL-18, as well as the secretion of other proteins such as IL-1α and fibroblast growth factor-2 through an unconventional protein secretion pathway[8]. Moreover, caspase-1 is required for pyroptosis, a form of cell death frequently observed during microbial infections that combines characteristics of apoptosis (DNA fragmentation) and necrosis (inflammation and cytokine release)[9]. Like other caspases, caspase-1 is synthesized as an inactive zymogen (pro-caspase-1) and becomes proteolytically active only after controlled dimerization in inflammasomes that are built around one of several different molecules[4,5,10–18] (Fig. 1). Whereas the leucine-rich repeat (LRR) domain is thought to be involved in autoinhibition that is disabled on direct or indirect sensing of the activating signal, the nucleotide-binding domain (NBD) is involved in the regulation of homo-oligomerization or hetero-oligomerization, which is required for inflammasome assembly. On receiving an activating signal, inflammasome sensors recruit pro-caspase-1 (which has a caspase activation and recruitment domain (CARD)) either directly through homotypic binding of CARD or indirectly through a pyrin domain (PYD) by means of the adaptor apoptosis-associated speck-like protein containing a CARD (ASC), which contains a PYD and a CARD.

### Diverse signals induce inflammasome formation

Inflammasome assembly is unique in its induction by a variety of both exogenous and endogenous signals. The range of activation signals sensed by each protein is distinct, but may include overlapping signals (Box 1). Whereas the AIM2 and NLRC4 inflammasomes are activated only by specific PAMPs, double-stranded DNA (dsDNA) and specific bacterial proteins, respectively[15–17], NLRP3 is activated by a large variety of signals, including PAMPs, DAMPs and bacterial toxins[19–22]. The structural diversity of ligands activating the NLRP3 inflammasome is in striking contrast to other innate pattern recognition receptors such as Toll-like receptors, which usually recognize more confined structural motifs[23]. Several models have therefore been proposed to explain how all of these signals could activate the NLRP3 inflammasome (Box 1). These non-exclusive mechanisms include both direct and indirect signal recognition mediated by additional accessory proteins[21,24,25]. We discuss the nature of some of these NLRP3-activating signals in the context of the pathophysiology of several diseases, but they have been more extensively reviewed elsewhere[4,5].

---

[1]Department of Immunobiology, Yale University, Connecticut 06520, USA. [2]Howard Hughes Medical Institute, Maryland 20815-6789, USA. *These authors contributed equally to this work.

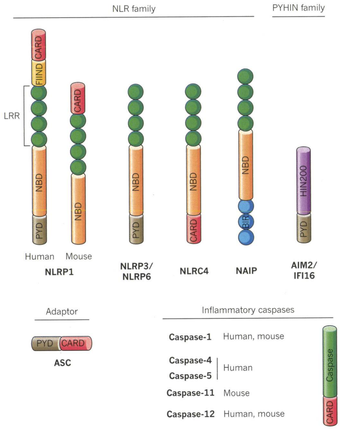


**Figure 1 | Domain organization of inflammasome proteins.** The identified core components belong to two families, the NOD-like receptor (NLR) family and the PYHIN (pyrin and HIN200 (haematopoietic interferon-inducible nuclear antigens with 200 amino-acid repeats) domain-containing protein) family. The NLR family members include NLRP1, NLRP2 (not shown), NLRP3, NLRP6, NLRC4 and, potentially, NLRP12. They all contain a nucleotide-binding domain (NBD), carboxy-terminal leucine-rich repeat (LRR), and can contain either a PYD or a caspase activation and recruitment domain (CARD) or both. The PYHIN family members AIM2 and IFI16 are characterized by having, in addition to a PYD, a HIN200 domain, which is involved in ligand binding. BIR, baculoviral inhibition of apoptosis repeat domain; FIIND, domain with function to find.

## Regulation of inflammasome activity

Inflammasome activity needs to be tightly regulated by the host to avoid the excess production of cytokines or overt cell death. Regulation occurs at transcriptional and post-transcriptional levels. First, the expression of inflammasome sensors, in particular NLRP3, is relatively low in many cell types and requires a priming signal to be induced[26]. This signal is frequently referred to as signal I and can be provided by microbial ligands, cytokines or reactive oxygen species (ROS). In addition, alternative splicing of inflammasome components generates protein variants with different activities. Indeed, splice variants of ASC have been identified with distinct abilities to serve as an inflammasome adaptor, with one variant even having inhibitory activity[27]. Furthermore, the host expresses proteins that regulate inflammasome activity mainly by sequestering inflammasome components through homotypic interaction of CARDs or PYDs or through directly inhibiting caspase-1 function[28,29]. Similarly, subcellular location and trafficking of inflammasome components seems to be important for the regulation of inflammasome activity because ASC is found mainly in the nucleus in resting cells[30]. Another level of regulation may occur through the assembly of complexes with different components, suggested by the observation that cytokine processing downstream of NLRC4 activation is ASC dependent, whereas pyroptosis is independent of ASC[31]. In addition, it has been suggested that differential processing of caspase-1 might contribute to this finding because catalytically active caspase-1 that has been rendered unable to perform autocleavage is impaired in its ability to cleave cytokines, but it is still able to initiate pyroptosis on NLRC4-mediated recognition of *Salmonella* infection[32]. Regulation of inflammasome activity is also achieved through crosstalk with cellular stress-associated processes such as autophagy. Induction of autophagy leads to the degradation of cellular substrates such as protein aggregates and organelles in autolysosomes for the recycling of metabolites. Strikingly, cells deficient in autophagy have a decreased threshold for inflammasome activation[33]. This has been suggested to be a result of the impaired clearance of defective mitochondria resulting in elevated levels of ROS, hinting at an involvement of NLRP3 as sensor[34,35]. Another aspect of the regulation of inflammasome activity is its downregulation either through secreted factors or cell–cell interactions. Examples of these signals are type I interferons or interactions between CD4$^+$ T cells and macrophages or dendritic cells, respectively, leading to transcriptional and post-transcriptional downregulation of inflammasome activity[36,37]. All of these process probably cooperate in the temporal and spatial organization of inflammasome-mediated processes. In summary, the host has evolved distinct mechanisms to regulate inflammasome activation to prevent the dire consequences of inflammasome overactivation.

## Regulation of inflammation by inflammasomes

Inflammasome-regulated processes depend on the simultaneous expression of the multiple inflammasome protein components in the same cell type of inflamed tissues. ASC and caspase-1 are found in many tissues and cell types, whereas the inflammasome sensors feature a more distinct expression pattern, suggesting tissue-specific mechanisms for sensing the microenvironment. In the following sections, we discuss how activation of inflammasomes in different lineages of cells regulates physiological reactions in the context of health and disease.

### Inflammasomes and the antimicrobial response

*In vivo*, inflammasomes have been shown to participate in the antimicrobial innate immune response[38]. The most widely studied inflammasome in this regard is the NLRP3 inflammasome, shown to be involved in antibacterial, viral, fungal and parasitic immune responses. Despite the evidence linking the NLRP3 inflammasome to the immune response to infection, only in a minority of cases has inflammasome activation by direct recognition of the pathogen been documented; many studies have indicated inflammasome activation through induction of signals related to cellular stress and damage[4] (see Box 1). The influenza A virus is an example of an indirect viral NLRP3 inflammasome activator. On infection, recognition of viral RNA by means of Toll-like receptor 7 (TLR7) induces transcription of the NLRP3 inflammasome components[39]. Subsequently, the activity of the viral ion-channel protein M2 induces pH neutralization of the *trans*-Golgi network, leading to potassium efflux and ROS formation, which in turn induce NLRP3 inflammasome assembly. Recently, messenger RNA of microbial origin was shown to activate the NLRP3 inflammasome in a TRIF (TIR-domain-containing adaptor-inducing interferon-β)-dependent manner, providing a mechanism for NLRP3 recognition of infection and the resultant differentiation by the host between viable and non-viable bacteria, leading to the induction of a potent immune response only on exposure to live microbes[40]. Mouse NLRP1 inflammasome activation is central to the initiation of the antimicrobial response to *Bacillus anthracis* infection through caspase-1-induced pyroptosis of infected macrophages, which permits self-limitation of infection and initiation of an antimicrobial neutrophilic reaction[41]. Human NLRP1 forms an ASC-dependent inflammasome, whereas mouse NLRP1b may activate caspase-1 in an ASC-independent manner[42]. NLRP1, like NOD2 and NLRP3, can sense muramyl dipeptide (MDP), a building block of the bacterial cell wall, suggesting possible cooperative roles of the different pattern recognition receptors in the sensing of and response to bacteria[43]. NLRC4 inflammasome activation is driven by type III and type IV secretion systems (T3SS and T4SS) of bacteria such as *Salmonella*, *Pseudomonas*, *Legionella* and *Yersinia*, which allow the cytoplasmic entry of the NLRC4 ligand flagellin, leading to activation of the NLRC4 pathway[31,44]. In addition, flagellin-independent activation of the NLRC4 inflammasome involves recognition of the T3SS rod protein PrgJ[45]. Interestingly, NLRC4 activation during *Legionella* infection is dependent on another NLR protein, NAIP5, but only partial dependence on NAIP5 was demonstrated for NLRC4 activation during *Pseudomonas* and *Salmonella* infection[45–47]. This apparent discrepancy was recently resolved by demonstrating that oligomerization of different NAIP family proteins with NLRC4 confers ligand specificity, so that a NAIP2–NLRC4 complex binds PrgJ and other structurally related rod proteins, whereas a NAIP5–NLRC4 complex binds flagellin[48]. Different pathogen receptors therefore assist NLRC4 in the recognition of microbial ligands, broadening the diversity of sensed structures. The recently discovered protein AIM2 recognizes bacterial and viral dsDNA, resulting in an antimicrobial response to intracellular pathogens, such as *Francisella tularensis*, *Listeria monocytogenes* and some DNA viruses[49,50].

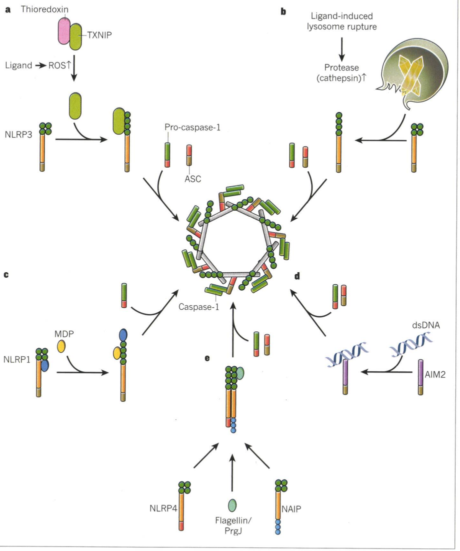
During microbial infection, distinct effector mechanisms of inflammasomes seem to be important for host defence. Although inflammasome-induced IL-1β and IL-18 are essential for the clearance of influenza virus and *Shigella*, respectively, pyroptosis has been suggested to be essential for caspase-1-mediated effects *in vivo* during infection by *Salmonella enterica* Typhimurium, *Legionella pneumophila* or *Burkholderia thailandensis*[51]. Altogether, it seems that the cooperative activity of distinct inflammasomes is pivotal in the host's ability to raise a protective immune response. This cooperation is illustrated during infection by *S. Typhimurium*, in which deficiencies in either NLRP3 or NLRC4 *in vivo* do not lead to increased bacterial infection, whereas mice deficient in both NLRC4 and NLRP3 have a similarly increased susceptibility to infection to that of caspase-1-deficient mice[52] (Fig. 2).

To avoid inducing inflammasome activation, many pathogens encode inhibitors of inflammasome function[53]. Frequently, these inhibitors, such as NS1 of influenza virus or Crm1 of cowpox virus, act on central inflammasome components and directly inhibit caspase-1 activity[54,55]. Other inhibitors, such as ORF63 of Kaposi's sarcoma-associated herpesvirus, bind NLR sensors and prevent the formation of NLRP1 and NLRP3 inflammasomes[56]. In general, the inhibition of inflammasome activity by the immune-evasive mechanisms of many microbes highlights the importance of these complexes in

**BOX 1**

## Models for inflammasome activation

Inflammasomes are assembled after sensing a structurally diverse repertoire of PAMPs and DAMPs. Several models have been proposed to explain how these signals are sensed, including models based on recognition of general cellular stress (Fig. **a** and **b**) or on direct and indirect recognition of activation signals (Fig. **c**–**e**). NLRP3 senses the reactive oxygen species (ROS), which is produced in the cell (potentially by mitochondria) directly or indirectly by activators of the NLRP3 inflammasome. Increased amounts of ROS are sensed by a complex of thioredoxin and thioredoxin-interacting protein (TXNIP), leading to the dissociation of this complex. Subsequent binding of TXNIP to NLRP3 leads to the activation of NLRP3, the recruitment of ASC and pro-caspase-1, and formation of the active inflammasome complex (**a**). NLRP3 is activated after lysosome destabilization. The phagocytosis of specific crystalline and particulate structures can lead to lysosome destabilization and the release of lysosomal content, including proteases. These proteases could lead to proteolytic inactivation of a negative regulator or to proteolytic activation of a positive regulator of NLRP3, resulting in inflammasome assembly (**b**). NLRP1 and AIM2 sense the ligand directly. The direct binding of specific ligands (muramyl dipeptide (MDP) and double-stranded DNA (dsDNA)) can lead to conformational changes in NLRP1 and AIM2, resulting in inflammasome activation. Inflammasome formation in NLRP1 is independent of ASC (**c, d**). NAIP proteins sense bacterial proteins resulting in the recruitment of NLRC4 and assembly of the NLRC4 inflammasome (**e**).



antimicrobial immunity. Over the next few years, it will be important to further characterize how pathogens are sensed by inflammasomes and how inflammasomes interact with additional innate immune pathways to regulate antimicrobial immune responses *in vivo*.

### Inflammasome activation by particulate compounds

In addition to infectious agents, there is evidence in mice and in humans that aberrant inflammasome activation by non-infectious agents may be linked to the pathogenesis of diseases characterized by sterile inflammation. For instance, crystal deposition diseases are a group of disorders in which inflammatory damage is elicited by exposure to exogenous or endogenous crystalline molecules. Examples of exogenous activators are silica and asbestos, whose endocytosis by pulmonary macrophages results in NLRP3 inflammasome activation involving ROS and lysosome destabilization, leading in turn to silicosis and asbestosis, respectively[24,57]. Similarly, aberrant formation of crystals from endogenous molecules such as monosodium urate (MSU) and calcium phosphate may lead to NLRP3 inflammasome activation in macrophages. Indeed, particular accumulation of MSU had long been noted during gout, but only recently has it been demonstrated that NLRP3-deficient mice feature defective MSU-induced neutrophil infiltration and inflammation[22] (Fig. 2). On the basis of these findings, clinical trials have been started to block IL-1β in gout; promising early results suggest that aberrant inflammasome activation drives inflammation in human crystal deposition disease. Furthermore, in patients with osteoarthritis, a common degenerative articular disorder, uric acid levels in synovial fluids are positively correlated with both IL-1β and IL-18 levels, as well as disease severity[58]. Calcium phosphate crystals were recently shown to activate NLRP3 (ref. 59). Hydroxyapatite crystals, a component of bone, are frequently found in osteoarthritis synovial fluid, activate IL-1β production by means of the NLRP3 inflammasome, and mediate inflammation and joint disease[59]. Future studies and clinical trials will be needed to explore whether targeting inflammasomes and their substrates are effective treatment options for these diseases.

An interesting example of inflammation that can be induced by an exogenous crystalline compound is the activation of the NLRP3 inflammasome by alum, an aluminium salt and the most commonly used adjuvant in human vaccines. Exposure of macrophages to alum *in vitro* leads to the NLRP3-dependent activation of caspase-1, and several studies have demonstrated that NLRP3 inflammasome-deficient mice have defects in alum-induced adaptive immune responses *in vivo*[60–62]. Several other studies have debated the role of the NLRP3 inflammasome in the adjuvant effect of alum, some suggesting that vaccination-induced adaptive immune responses may be ASC dependent but independent of caspase-1 (and hence independent of inflammasomes)[63–66]. A recent study suggested an alternative NLRP3-independent mechanism for alum's adjuvant effects, namely attachment to the plasma membrane of dendritic cells, leading to abortive phagocytosis and imprinting dendritic cells to induce a humoral immune response[67]. The precise mechanisms related to vaccination adjuvant effects and the possible contribution by inflammasomes remain to be clarified in future studies.

## Inflammasomes and metabolic syndrome

The incidence of obesity worldwide has increased markedly in recent decades. Obesity is associated with multiorgan (namely pancreatic, adipose, hepatic, cardiac and muscle tissue) chronic metabolic and inflammatory alterations that together are termed 'metabolic syndrome'. These alterations include impaired insulin sensitivity, pancreatic β-cell dysfunction, non-alcoholic fatty liver disease and atherosclerosis. Unequivocal experimental and clinical evidence causally link IL-1β and IL-18 to the development of these metabolic pathologies and their complications[68,69]. However, the underlying molecular mechanisms that result in tissue-specific inflammasome activation in the context of obesity have only just begun to be explained.

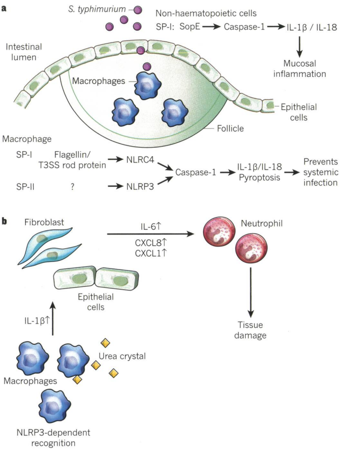


**Figure 2 | Inflammasome activity regulates inflammation during microbial infection and autoinflammatory diseases.** **a,** *S. typhimurium* enters the host by crossing through the intestinal epithelial barrier. M cells, a specialized epithelial cell type layering Peyer's patches, are particularly involved in the transcytosis of *Salmonella* and the infection of macrophages in Peyer's patches. The inflammasome and caspase-1 are involved in several cell types and at several steps of the infection. Injection of the bacterial effector protein SopE into epithelial cells induces the activation of caspase-1 independently of NLRP3 and NLRC4 through a process involving the GTPase Rac1. The resultant mucosal inflammation is dependent on IL-1β and IL-18 produced by non-haematopoietic cells. On infection of macrophages, the bacterial proteins flagellin and PrgJ (part of the T3SS) are sensed by means of NLRC4. This results in the activation of caspase-1, leading to both IL-1β/IL-18 processing and pyroptosis, which limit systemic infection. NLRP3 contributes to these processes by recognition of an unknown signal. **b,** Phagocytosis of monosodium urate (MSU) crystals by macrophages induces the NLRP3-dependent caspase-1 activation and release of IL-1β, which stimulates non-haematopoietic cells to produce IL-6 and chemokines (CXCL1 and CXCL8), attracting neutrophils. Activated neutrophils then cause tissue damage. Therapeutic blockade of IL-1β in humans ameliorates inflammatory bouts in gout.

### Inflammasome activation regulates insulin signalling

NLRP3 inflammasome components and caspase-1 activation are increased in the adipose tissue and liver of obese mice and humans; moreover, their level of expression is directly correlated with the severity of type 2 diabetes mellitus (T2DM) in obese individuals[70]. During nutritional surplus, in addition to adipocyte hypertrophy owing to increased lipid storage, adipose tissue is infiltrated by classically activated, M1, macrophages that secrete pro-inflammatory cytokines[71]. NLRP3, ASC and caspase-1 are preferentially expressed in adipose-tissue-infiltrating macrophages, in which the saturated fatty acid palmitate and lipotoxic ceramides trigger NLRP3 inflammasome activation through a mechanism that involves defective autophagy and the accumulation of mitochondrial ROS[70,72] (Fig. 3). Enhanced

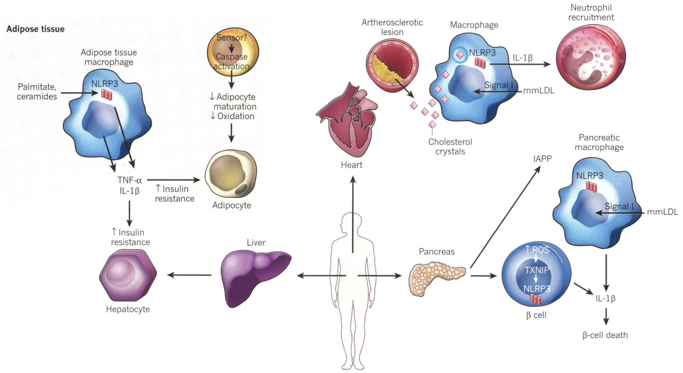


**Figure 3 | The role of inflammasomes in metabolic syndrome.** During obesity, the NLRP3 inflammasome is activated by obesity-associated DAMPs in multiple tissues and cell types; the resultant pro-inflammatory-induced state often leads to a deterioration in metabolic functions. In adipose tissue, palmitate and ceramides activate the NLRP3 inflammasome in infiltrating macrophages, which leads to an enhancement of insulin resistance. In addition, caspase-1 activation through an unknown sensor protein regulates adipocyte differentiation and fatty acid oxidation. In the pancreas, IAPP and increased mitochondrial ROS production activate the NLRP3 inflammasome in mmLDL-primed macrophages and β cells, respectively. The increased levels of IL-1β in pancreatic islets result in increased β-cell death and decreased insulin production. Minute cholesterol crystals in early atherosclerotic lesions activate the NLRP3 inflammasome in mmLDL-primed macrophages, promoting inflammatory-cell infiltration and increased atherosclerosis progression.

NLRP3-mediated caspase-1 activation in adipose-tissue-infiltrating macrophages of obese mice results in at least two deleterious consequences in insulin-sensitive tissues: first, IL-1β inhibits insulin signalling by direct serine phosphorylation of insulin receptor substrate 1 and induces the expression of tumour necrosis-factor-α, a well-characterized insulin-resistance-promoting cytokine; and second, IL-1β and IL-18 induce type 1 CD4 T helper cells in adipose tissue[70].

Inflammasomes also seem to have adipocyte-intrinsic functions. Caspase-1 is upregulated during adipocyte differentiation, in which it promotes insulin resistance through autocrine IL-1β-mediated effects; however, perhaps more interestingly, caspase-1-deficient precursors differentiate more efficiently into mature adipocytes and have an increased oxidation rate[73] (Fig. 3), which suggests that adipocyte-specific proteins may serve as caspase-1 substrates, leading to functional changes in adipose tissue of obese animals. Hence, *Nlrp3*$^{-/-}$, *Asc*$^{-/-}$ and *Casp1*$^{-/-}$ mice have been reported to have decreased weight gain and fat mass, as well as decreased insulin resistance. Interestingly, the loss of function of NLRP3 decreases but does not eliminate caspase-1 activation in adipose tissue or liver, suggesting that additional inflammasome sensors might contribute to the pathophysiology of obesity; furthermore, the nature of the stimuli that trigger caspase-1 activation during the accumulation of fat in adipocytes remains to be determined.

## Inflammasome activation impairs β-cell function

Chronic hyperglycaemia as a result of peripheral insulin resistance is compensated for by increased insulin output by pancreatic β cells. Local inflammatory processes coupled with the toxic effects of glucose lead to accelerated mass loss of β cells and decreased insulin secretion over time, which prompts the progression from obesity and insulin resistance to overt T2DM. IL-1β, preferentially expressed by pancreatic infiltrating macrophages and to a smaller extent by β cells, has been implicated as a critical driver of β-cell death in conditions of chronic exposure to elevated concentrations of glucose[74]. Additional support for a pathological role of inflammasomes in T2DM comes from a recent report showing that glyburide, an insulin-secretion-promoting drug, suppresses NLRP3-mediated IL-1β release[75], and more importantly from human clinical trials of IL-1 receptor antagonist (IL-1RA) to treat T2DM that demonstrated improved glycaemic control and β-cell function[76].

In pancreatic islet failure during T2DM, NLRP3 inflammasomes may be activated by different mechanisms. First, hyperglycaemia triggers mitochondrial ROS production in β cells by increasing the activity of the electron transport chain, which induces thioredoxin-interacting protein (TXNIP) to dissociate from thioredoxin, leading to the activation of NLRP3 (ref. 25) (Fig. 3 and Box 1). Second, the severity of T2DM is closely correlated with the levels of deposition of amyloid polypeptide (IAPP, also known as amylin) in pancreatic islets[77]. Remarkably, IAPP has been shown to specifically activate the NLRP3 inflammasome in mouse pancreatic macrophages previously primed with the abundant minimally oxidized low-density lipoprotein (mmLDL) through a pathway that involves disruption of the phagolysosomal pathway[78] (Fig. 3). Further evidence for the role of IAPP and NLRP3 activation in pancreatic islet deterioration is provided by the observation that macrophages express more IL-1β in pancreatic islets of human *IAPP*-transgenic mice[78]. Overall, secreted IL-1β can then signal in an autocrine or paracrine manner to induce β-cell death or dysfunction and to promote the expression of chemotactic factors that further worsen immune-cell infiltration. It is important to mention that TXNIP is not required for NLRP3 activation in macrophages, strengthening the notion that the cellular milieu is a critical determinant of the way in which the NLRP3 inflammasome senses damage signals in the context of metabolic abnormalities[78]. However, despite great advances in recent years, several questions remain unresolved, including the mechanisms of pro-IL-1β induction in β cells, the exact contribution of inflammasome activation in infiltrating macrophages and β cells to disease progression, and the potential roles of additional inflammasome sensors in pancreatic inflammatory processes.

### Inflammasome activation and atherosclerosis

The inflammatory nature of atherosclerosis is well established, but the biological agents that trigger artery wall inflammation remain largely unknown. Cholesterol crystal deposition in arterial vessels has long been a pathognomonic feature of atherosclerosis; moreover, recent evidence suggests that they are present at early stages of atherosclerotic lesions, coinciding with the first appearance of inflammatory cells[79]. Similarly to IAPP, cholesterol crystals activate the NLRP3 inflammasome through phagolysosome destabilization in mmLDL-primed mouse and human macrophages[79] (Fig. 3). LDL-receptor-deficient mice (prone to atherosclerosis) reconstituted with bone marrow deficient in NLRP3, ASC or IL-1α/β are markedly resistant to the development of atherosclerosis, suggesting that NLRP3 inflammasome activation and IL-1 secretion from the haematopoietic compartment are key events in the early stages of disease[79]. Another model for atherosclerosis is the *Apoe*$^{-/-}$ mouse, which develops severe hypercholesterolaemia and spontaneous atherosclerosis when being fed with a high-fat diet[80]. Earlier reports have demonstrated that IL-1-receptor-deficient *Apoe*$^{-/-}$ mice and IL-1RA-treated *Apoe*$^{-/-}$ mice showed decreased atherosclerosis, suggesting a role for the inflammasome in atherogenesis in this mouse model[81,82]. Interestingly, using mice deficient in both ApoE and different components of the NLRP3 inflammasome, Menu *et al.*[83] showed that atherosclerosis progresses independently of the NLRP3 inflammasome in this context. The most reasonable explanations for the discrepancies between the two studies lie in the differences between atherosclerosis models (*Apoe*$^{-/-}$ versus *Ldlr*$^{-/-}$) and in a putative role for IL-1α in the *Apoe*$^{-/-}$ mouse model. However, it also raises the interesting possibility that environmental factors could account for the contradictory phenotypes. Clearly, further research is warranted regarding the implication of the inflammasomes in the pathogenesis of atherosclerosis and their role in non-haematopoietic cells during atherosclerotic plaque formation.

## Inflammasomes and the mucosal immune response

In the gastrointestinal mucosa, the host is separated from an immense microbial ecosystem by only a single layer of epithelial cells[84]. To avoid ongoing inflammatory reactions to commensal microbes and food antigens, while preserving the ability to react to pathogenic insults, mammals have evolved a complex mucosal immune system composed of epithelial and stromal cells acting together with subsets of cells of haematopoietic origin. These cells interact closely with each other and with the surrounding microbial milieu. When such interactions are perturbed, autoinflammation may develop, potentially leading to IBD. The ability of inflammasomes to recognize exogenous and endogenous signals has led to several studies characterizing their role during chemical-induced intestinal autoinflammation, a model for human IBD. Interestingly, conflicting results have been reported by several groups (Fig. 4). Using a common acute and chronic epithelial injury colitis model based on the administration of dextran sulphate to mice, several groups reported decreased disease severity in mice deficient in caspase-1 or NLRP3, which correlated with lower IL-1β production during disease[85,86]. Using the same model, other groups found that mice deficient in NLRP3, ASC and caspase-1 show an exacerbated disease severity[87–90]. In these reports, a role was suggested for the NLRP3 inflammasome in the promotion of tissue regeneration in response to injury. Although variations in these results could be explained by differences in experimental designs, a recent paper[91] has offered an alternative explanation. This study reported that an NLRP6 inflammasome participates in the steady-state regulation of the commensal microflora. Deficiency in this inflammasome is associated with an alteration in the microflora communities and the emergence of normally suppressed bacteria that have pro-inflammatory activity. It was suggested that regulation of the microflora by epithelial inflammasomes is mediated, at least in part, through the induction of basal secretion of IL-18 by epithelial cells. Discrepancies in the earlier studies may therefore be explained by differences in microflora communities in different animal facilities, coupled with the inability of NLRP6-inflammasome-deficient mice (*Nlrp6*$^{-/-}$, *Asc*$^{-/-}$, *Casp1*$^{-/-}$) to regulate the microflora in a given facility.

Phenotypic alterations may be induced by differences in animal housing; wild-type mice housed with NLRP6-inflammasome-deficient mice acquire their colitogenic microflora and develop phenotypes that may be profoundly different from that of wild-type mice housed alone. The factors inducing the formation of the NLRP6 inflammasome and the precise effector mechanisms for regulation of the microflora remain to be studied. In addition to NLRP6, other inflammasomes that are expressed mainly within the haematopoietic compartment, such as the NLRP3 inflammasome, may also function by regulating the microflora, as well as the autoinflammatory process itself. In the mucosal environment, inflammasomes therefore function in the sensing of pathogens and the commensal flora by non-haematopoietic cells, such as the epithelial layers, but also by haematopoietic cells. They cooperate to maintain tolerance towards commensal microbes and to initiate a potent immune response towards pathogens (Fig. 4). Distinct inflammasomes expressed in different cell lineages may orchestrate these seemingly opposite functions during acute mucosal inflammation.

It has been suggested that inflammasomes participate in inflammation-induced tumorigenesis, a common complication of chronic intestinal autoinflammation (such as IBD), yet mechanisms accounting for this regulatory effect remain elusive[87,89]. Two recent reports suggest that NLRP6 deficiency may be associated with enhanced inflammation-induced tumorigenesis through impaired mucosal self-renewal and proliferation mediated by alterations in the intestinal stem-cell niche, or alternatively by an impaired haematopoietic immune response[92,93]. It remains to be determined whether the effect of NLRP6 on tumorigenesis is a direct consequence of NLRP6 deficiency in one or more intestinal cell types or whether it is indirectly induced by the documented alterations in the regulation of microflora composition of these mice.

Alternatively, another report suggests that epithelial NLRC4 inflammasome activity may also participate in the prevention of intestinal inflammation-induced tumorigenesis, because NLRC4-deficient mice were more prone to the development of colonic neoplasms after the induction of chronic inflammation[94]. This effect was independent of inflammation or the composition of the intestinal microflora and was suggested to involve enhanced proliferation of epithelial cells and impaired apoptosis of colonic epithelial cells. The precise molecular mechanism linking the NLRC4 inflammasome dysfunction to increased tumorigenesis remains to be explored.

## Future perspectives

The discovery of inflammasomes as protein platforms that control the processing of IL-1β and IL-18 constitutes a milestone in the innate immunology field. However, many questions remain unanswered, including those regarding the biochemical and genetic regulation of these protein complexes, as well as their roles in complex diseases. These, in our opinion, may be the focus of inflammasome research in coming years.

The mechanism through which sensor proteins recognize activating signals remains a critical unresolved issue in the field. Although *in vitro* reconstitution assays with purified components have been performed with NLRP1 and AIM2 and have shown direct binding of the respective ligands to both sensors, intense research has not been successful in identifying such direct interactions with any of the other inflammasome sensors[15–17,43]. Hence, it is plausible that as yet unidentified proteins may facilitate the recognition of structurally diverse activating

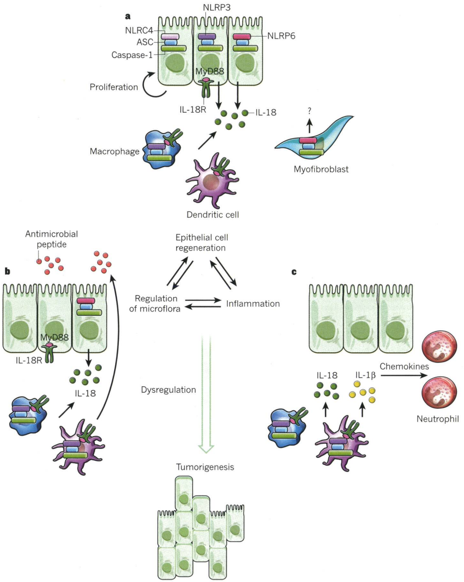


**Figure 4 | Inflammasomes regulate multiple aspects of tissue homeostasis and immune response in the mucosa.** NLRP3, NLRP6 and NLRC4 inflammasomes are formed in many lineages of haematopoietic and non-haematopoietic cells present in the intestine. The activity of these inflammasomes regulates epithelial cell regeneration, the microflora and mucosal inflammation. All of these processes are interconnected, and the dysregulation of any of them may increase the risk of tumorigenesis. **a**, Epithelial cell regeneration: recognition of as yet unknown signals leads to the processing of pro-IL-18 and potentially unidentified effectors through NLRP3, NLRP6 and NLRC4. This allows the regulation of epithelial regeneration on injury, through effects on epithelial cells, lamina propria immune cells, myofibroblasts and intestinal epithelial stem cells. **b**, Regulation of microflora: during the steady state, the production of IL-18 through NLRP3 and NLRP6 in the colon participates in the regulation of a 'healthy' microflora, potentially through regulation of antimicrobial peptide secretion by epithelial cells and immune cell subsets. Disruption of the NLRP6 inflammasome results in aberrant regulation of the intestinal microflora, leading to the emergence of pro-colitogenic bacteria. **c**, Inflammation: after breaching of the intestinal epithelial barrier, the NLRP3 inflammasome is involved in the regulation of intestinal inflammation in response to microbial invasion. Both IL-1β and IL-18 are suggested to act as effector molecules, leading to the recruitment of multiple cell subsets, including neutrophils, that promote local inflammation.

signals and ligands by NLRP3. Such 'bridging' mechanisms have been suggested to link cellular stress (namely potassium efflux, generation of ROS and lysosome destabilization) and microbial proteins to NLRP3 (refs 21, 24, 25). However, the exact identity of these bridging molecules remains to be fully characterized. Indeed, for the NLRC4 inflammasomes, NAIP proteins were identified as additional components that are responsible for the specificity of microbial ligands recognized by the complexes[48]. Another open issue is the interesting observation that several phenotypes related to a deficiency of common inflammasome effectors and adaptors cannot be attributed to the well-known NLR and PYHIN molecular sensors and their ligands, indicating the existence of additional sensors and ligands to be identified. An example of a recently identified sensor is IFI16, a sensor of viral DNA[18].

The early discovery that mutations in inflammasome components in humans are the cause of several inherited autoinflammatory diseases that can be treated by neutralizing IL-1β illustrates the potential detrimental effects of overactivation of inflammasomes[95]. However, recent evidence from mouse studies on the role of inflammasomes in mucosal immunology indicates a compartmentalized and dual function (detrimental versus beneficial). To dissect the contributions of inflammasomes in different cell types and tissues, novel tools such as conditional gene-knockout mice will have to be developed. This would allow better investigation of areas that bear fundamental importance and present complex challenges to the research community, such as the involvement of inflammasome signalling in circuits regulating metabolism, inflammation and the intestinal microflora in mice and humans[96].

It is worth noting that NLRP6 inflammasome deficiency was recently shown to associate with expansion of an autoinflammation-promoting microbiota including bacterial taxa such as Prevotellaceae and TM7 (ref. 91). The mechanisms by which inflammasomes sense microflora members or communities, regulate tissue repair and regeneration, and orchestrate mucosal immune responses during the steady state and inflammation remain to be fully characterized. Interestingly, both inflammasome deficiency and alterations in the intestinal microflora were previously linked in separate studies to a propensity for the development of elements of the metabolic syndrome, such as obesity and atherosclerosis, in mice and humans[97,98]. Indeed, recent characterization of the microflora community in a cohort of individuals of European origin identified the bacterial family Prevotellaceae as a prominent component of one human microflora cluster that featured a trend towards increased body mass index[99]. Moreover, faecal contents of rural African children featured expanded representation of Prevotellaceae that was absent in their European counterparts, probably reflecting a combination of genetic and environmental factors that differentially regulate the microflora composition[100]. Whether the regulation of the gut flora by inflammasomes affects weight, metabolism and inflammation is expected to become a major interest of the field.

Whereas the first decade of inflammasome research has been characterized mostly by *in vitro* studies and *in vivo* studies with small animal models, the contribution of altered inflammasome function to complex human disease is expected to gradually take centre stage. This, together with a vast effort to discover and develop small-molecule approaches to the specific inhibition of inflammasome activation, may pave the way to therapeutic intervention targeting inflammasome-regulated pathways that are involved in the pathogenesis of human disease. ■

1. Medzhitov, R. Origin and physiological roles of inflammation. *Nature* **454,** 428–435 (2008).
2. Medzhitov, R. & Janeway, C. A. Jr. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* **91,** 295–298 (1997).
3. Matzinger, P. Tolerance, danger, and the extended family. *Annu. Rev. Immunol.* **12,** 991–1045 (1994).
4. Schroder, K. & Tschopp, J. The inflammasomes. *Cell* **140,** 821–832 (2010).
5. Davis, B. K., Wen, H. & Ting, J. P. The inflammasome NLRs in immunity, inflammation, and associated diseases. *Annu. Rev. Immunol.* **29,** 707–735 (2011).
6. Dinarello, C. A. IL-1: discoveries, controversies and future directions. *Eur. J. Immunol.* **40,** 599–606 (2010).
7. Cerretti, D. P. *et al.* Molecular cloning of the interleukin-1β converting enzyme. *Science* **256,** 97–100 (1992).
8. Keller, M., Ruegg, A., Werner, S. & Beer, H. D. Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132,** 818–831 (2008).
9. Bergsbaken, T., Fink, S. L. & Cookson, B. T. Pyroptosis: host cell death and inflammation. *Nature Rev. Microbiol.* **7,** 99–109 (2009).
10. Agostini, L. *et al.* NALP3 forms an IL-1β-processing inflammasome with increased activity in Muckle–Wells autoinflammatory disorder. *Immunity* **20,** 319–325 (2004).
11. Martinon, F., Burns, K. & Tschopp, J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-β. *Mol. Cell* **10,** 417–426 (2002).
12. Grenier, J. M. *et al.* Functional screening of five PYPAF family members identifies PYPAF5 as a novel regulator of NF-κB and caspase-1. *FEBS Lett.* **530,** 73–78 (2002).
13. Wang, L. *et al.* PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates activation of NF-κB and caspase-1-dependent cytokine processing. *J. Biol. Chem.* **277,** 29874–29880 (2002).
14. Poyet, J. L. *et al.* Identification of Ipaf, a human caspase-1-activating protein related to Apaf-1. *J. Biol. Chem.* **276,** 28309–28313 (2001).
15. Fernandes-Alnemri, T., Yu, J. W., Datta, P., Wu, J. & Alnemri, E. S. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458,** 509–513 (2009).
16. Hornung, V. *et al.* AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458,** 514–518 (2009).
17. Roberts, T. L. *et al.* HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA. *Science* **323,** 1057–1060 (2009).
18. Kerur, N. *et al.* IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi sarcoma-associated herpesvirus infection. *Cell Host Microbe* **9,** 363–375 (2011).
19. Mariathasan, S. *et al.* Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature* **440,** 228–232 (2006).
20. Dostert, C. *et al.* Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica. *Science* **320,** 674–677 (2008).
21. Petrilli, V. *et al.* Activation of the NALP3 inflammasome is triggered by low intracellular potassium concentration. *Cell Death Differ.* **14,** 1583–1589 (2007).
22. Martinon, F., Petrilli, V., Mayor, A., Tardivel, A. & Tschopp, J. Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* **440,** 237–241 (2006).
23. Jin, M. S. & Lee, J. O. Structures of the Toll-like receptor family and its ligand complexes. *Immunity* **29,** 182–191 (2008).
24. Hornung, V. *et al.* Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nature Immunol.* **9,** 847–856 (2008).
25. Zhou, R., Tardivel, A., Thorens, B., Choi, I. & Tschopp, J. Thioredoxin-interacting protein links oxidative stress to inflammasome activation. *Nature Immunol.* **11,** 136–140 (2010).
26. Hornung, V. & Latz, E. Critical functions of priming and lysosomal damage for NLRP3 activation. *Eur. J. Immunol.* **40,** 620–623 (2010).
27. Bryan, N. B. *et al.* Differential splicing of the apoptosis-associated speck like protein containing a caspase recruitment domain (ASC) regulates inflammasomes. *J. Inflamm.* **7,** 23 (2010).
28. Stehlik, C. & Dorfleutner, A. COPs and POPs: modulators of inflammasome activity. *J. Immunol.* **179,** 7993–7998 (2007).
29. Young, J. L. *et al.* The serpin proteinase inhibitor 9 is an endogenous inhibitor of interleukin 1β-converting enzyme (caspase-1) activity in human vascular smooth muscle cells. *J. Exp. Med.* **191,** 1535–1544 (2000).
30. Bryan, N. B., Dorfleutner, A., Rojanasakul, Y. & Stehlik, C. Activation of inflammasomes requires intracellular redistribution of the apoptotic speck-like protein containing a caspase recruitment domain. *J. Immunol.* **182,** 3173–3182 (2009).
31. Mariathasan, S. *et al.* Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. *Nature* **430,** 213–218 (2004).
32. Broz, P., von Moltke, J., Jones, J. W., Vance, R. E. & Monack, D. M. Differential requirement for Caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell Host Microbe* **8,** 471–483 (2010).
33. Saitoh, T. *et al.* Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1β production. *Nature* **456,** 264–268 (2008).
34. Zhou, R., Yazdi, A. S., Menu, P. & Tschopp, J. A role for mitochondria in NLRP3 inflammasome activation. *Nature* **469,** 221–225 (2011).
35. Nakahira, K. *et al.* Autophagy proteins regulate innate immune responses by inhibiting the release of mitochondrial DNA mediated by the NALP3 inflammasome. *Nature Immunol.* **12,** 222–230 (2011).
36. Guarda, G. *et al.* Type I interferon inhibits interleukin-1 production and inflammasome activation. *Immunity* **34,** 213–223 (2011).
37. Guarda, G. *et al.* T cells dampen innate immune responses through inhibition of NLRP1 and NLRP3 inflammasomes. *Nature* **460,** 269–273 (2009).
38. Elinav, E., Strowig, T., Henao-Mejia, J. & Flavell, R. A. Regulation of the antimicrobial response by NLR proteins. *Immunity* **34,** 665–679 (2011).
39. Ichinohe, T., Pang, I. K. & Iwasaki, A. Influenza virus activates inflammasomes via its intracellular M2 ion channel. *Nature Immunol.* **11,** 404–410 (2010).
40. Sander, L. E. *et al.* Detection of prokaryotic mRNA signifies microbial viability and promotes immunity. *Nature* **474,** 385–389 (2011).
41. Boyden, E. D. & Dietrich, W. F. *Nalp1b* controls mouse macrophage susceptibility to anthrax lethal toxin. *Nature Genet.* **38,** 240–244 (2006).
42. Hsu, L. C. *et al.* A NOD2–NALP1 complex mediates caspase-1-dependent IL-1β secretion in response to *Bacillus anthracis* infection and muramyl dipeptide. *Proc. Natl Acad. Sci. USA* **105,** 7803–7808 (2008).
43. Faustin, B. *et al.* Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Mol. Cell* **25,** 713–724 (2007).
44. Sutterwala, F. S. *et al.* Immune recognition of *Pseudomonas aeruginosa* mediated by the IPAF/NLRC4 inflammasome. *J. Exp. Med.* **204,** 3235–3245 (2007).
45. Miao, E. A. *et al.* Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. *Proc. Natl Acad. Sci. USA* **107,** 3076–3080 (2010).
46. Lightfield, K. L. *et al.* Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. *Nature Immunol.* **9,** 1171–1178 (2008).
47. Lamkanfi, M. *et al.* The Nod-like receptor family member Naip5/Birc1e restricts *Legionella pneumophila* growth independently of caspase-1 activation. *J. Immunol.* **178,** 8022–8027 (2007).
48. Kofoed, E. M. & Vance, R. E. Innate immune recognition of bacterial ligands by NAIPs determines inflammasome specificity. *Nature* **477,** 592–595 (2011).
49. Fernandes-Alnemri, T. *et al.* The AIM2 inflammasome is critical for innate immunity to *Francisella tularensis*. *Nature Immunol.* **11,** 385–393 (2010).
50. Rathinam, V. A. *et al.* The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. *Nature Immunol.* **11,** 395–402 (2010).
51. Miao, E. A. *et al.* Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria. *Nature Immunol.* **11,** 1136–1142 (2010).
52. Broz, P. *et al.* Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against *Salmonella*. *J. Exp. Med.* **207,** 1745–1755 (2010).
53. Taxman, D. J., Huang, M. T. & Ting, J. P. Inflammasome inhibition as a pathogenic stealth mechanism. *Cell Host Microbe* **8,** 7–11 (2010).
54. Stasakova, J. *et al.* Influenza A mutant viruses with altered NS1 protein function provoke caspase-1 activation in primary human macrophages, resulting in fast apoptosis and release of high levels of interleukins 1β and 18. *J. Gen. Virol.* **86,** 185–195 (2005).
55. Ray, C. A. *et al.* Viral inhibition of inflammation: cowpox virus encodes an inhibitor of the interleukin-1β converting enzyme. *Cell* **69,** 597–604 (1992).
56. Gregory, S. M. *et al.* Discovery of a viral NLR homolog that inhibits the inflammasome. *Science* **331,** 330–334 (2011).
57. Cassel, S. L. *et al.* The Nalp3 inflammasome is essential for the development of silicosis. *Proc. Natl Acad. Sci. USA* **105,** 9035–9040 (2008).
58. Denoble, A. E. *et al.* Uric acid is a danger signal of increasing risk for osteoarthritis through inflammasome activation. *Proc. Natl Acad. Sci. USA* **108,** 2088–2093 (2011).
59. Pazar, B. *et al.* Basic calcium phosphate crystals induce monocyte/macrophage IL-1β secretion through the NLRP3 inflammasome *in vitro*. *J. Immunol.* **186,** 2495–2502 (2011).
60. Kool, M. *et al.* Alum adjuvant stimulates inflammatory dendritic cells through activation of the NALP3 inflammasome. *J. Immunol.* **181,** 3755–3759 (2008).
61. Li, H., Willingham, S. B., Ting, J. P. & Re, F. Inflammasome activation by alum and alum's adjuvant effect are mediated by NLRP3. *J. Immunol.* **181,** 17–21 (2008).
62. Eisenbarth, S. C., Colegio, O. R., O'Connor, W., Sutterwala, F. S. & Flavell, R. A. Crucial role for the Nalp3 inflammasome in the immunostimulatory properties of aluminium adjuvants. *Nature* **453,** 1122–1126 (2008).
63. Kuroda, E. *et al.* Silica crystals and aluminum salts regulate the production of prostaglandin in macrophages via NALP3 inflammasome-independent mechanisms. *Immunity* **34,** 514–526 (2011).
64. Ellebedy, A. H. *et al.* Inflammasome-independent role of the apoptosis-associated speck-like protein containing CARD (ASC) in the adjuvant effect of MF59. *Proc. Natl Acad. Sci. USA* **108,** 2927–2932 (2011).
65. McKee, A. S. *et al.* Alum induces innate immune responses through macrophage and mast cell sensors, but these sensors are not required for alum to act as an adjuvant for specific immunity. *J. Immunol.* **183,** 4403–4414 (2009).

66. Franchi, L. & Nunez, G. The Nlrp3 inflammasome is critical for aluminium hydroxide-mediated IL-1β secretion but dispensable for adjuvant activity. *Eur. J. Immunol.* **38,** 2085–2089 (2008).
67. Flach, T. L. *et al.* Alum interaction with dendritic cell membrane lipids is essential for its adjuvanticity. *Nature Med.* **17,** 479–487 (2011).
68. Feve, B. & Bastard, J. P. The role of interleukins in insulin resistance and type 2 diabetes mellitus. *Nature Rev. Endocrinol.* **5,** 305–311 (2009).
69. Netea, M. G. *et al.* Deficiency of interleukin-18 in mice leads to hyperphagia, obesity and insulin resistance. *Nature Med.* **12,** 650–656 (2006).
70. Vandanmagsar, B. *et al.* The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. *Nature Med.* **17,** 179–188 (2011).
71. Lumeng, C. N., Bodzin, J. L. & Saltiel, A. R. Obesity induces a phenotypic switch in adipose tissue macrophage polarization. *J. Clin. Invest.* **117,** 175–184 (2007).
72. Wen, H. *et al.* Fatty acid-induced NLRP3-ASC inflammasome activation interferes with insulin signaling. *Nature Immunol.* **12,** 408–415 (2011).
73. Stienstra, R. *et al.* The inflammasome-mediated caspase-1 activation controls adipocyte differentiation and insulin sensitivity. *Cell Metab.* **12,** 593–605 (2010).
74. Dinarello, C. A., Donath, M. Y. & Mandrup-Poulsen, T. Role of IL-1β in type 2 diabetes. *Curr. Opin. Endocrinol. Diabetes Obes.* **17,** 314–321 (2010).
75. Lamkanfi, M. *et al.* Glyburide inhibits the cryopyrin/Nalp3 inflammasome. *J. Cell Biol.* **187,** 61–70 (2009).
76. Larsen, C. M. *et al.* Interleukin-1-receptor antagonist in type 2 diabetes mellitus. *N. Engl. J. Med.* **356,** 1517–1526 (2007).
77. Seino, S & Study Group of Comprehensive Analysis of Genetic Factors in Diabetes Mellitus. S20G mutation of the amylin gene is associated with type II diabetes in Japanese. Study Group of Comprehensive Analysis of Genetic Factors in Diabetes Mellitus. *Diabetologia* **44,** 906–909 (2001).
78. Masters, S. L. *et al.* Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1β in type 2 diabetes. *Nature Immunol.* **11,** 897–904 (2010).
79. Duewell, P. *et al.* NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature* **464,** 1357–1361 (2010).
80. Plump, A. S. *et al.* Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell* **71,** 343–353 (1992).
81. Chi, H., Messas, E., Levine, R. A., Graves, D. T. & Amar, S. Interleukin-1 receptor signaling mediates atherosclerosis associated with bacterial exposure and/or a high-fat diet in a murine apolipoprotein E heterozygote model: pharmacotherapeutic implications. *Circulation* **110,** 1678–1685 (2004).
82. Isoda, K. & Ohsuzu, F. The effect of interleukin-1 receptor antagonist on arteries and cholesterol metabolism. *J. Atheroscler. Thromb.* **13,** 21–30 (2006).
83. Menu, P. *et al.* Atherosclerosis in ApoE-deficient mice progresses independently of the NLRP3 inflammasome. *Cell Death Dis.* **2,** e137 (2011).
84. Maloy, K. J. & Powrie, F. Intestinal homeostasis and its breakdown in inflammatory bowel disease. *Nature* **474,** 298–306 (2011).
85. Siegmund, B., Lehr, H. A., Fantuzzi, G. & Dinarello, C. A. IL-1β -converting enzyme (caspase-1) in intestinal inflammation. *Proc. Natl Acad. Sci. USA* **98,** 13249–13254 (2001).
86. Bauer, C. *et al.* Colitis induced in mice with dextran sulfate sodium (DSS) is mediated by the NLRP3 inflammasome. *Gut* **59,** 1192–1199 (2010).
87. Allen, I. C. *et al.* The NLRP3 inflammasome functions as a negative regulator of tumorigenesis during colitis-associated cancer. *J. Exp. Med.* **207,** 1045–1056 (2010).
88. Hirota, S. A. *et al.* NLRP3 inflammasome plays a key role in the regulation of intestinal homeostasis. *Inflamm. Bowel Dis.* **17,** 1359–1372 (2011).
89. Zaki, M. H. *et al.* The NLRP3 inflammasome protects against loss of epithelial integrity and mortality during experimental colitis. *Immunity* **32,** 379–391 (2010).
90. Dupaul-Chicoine, J. *et al.* Control of intestinal homeostasis, colitis, and colitis-associated colorectal cancer by the inflammatory caspases. *Immunity* **32,** 367–378 (2010).
91. Elinav, E. *et al.* NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* **145,** 745–757 (2011).
92. Normand, S. *et al.* Nod-like receptor pyrin domain-containing protein 6 (NLRP6) controls epithelial self-renewal and colorectal carcinogenesis upon injury. *Proc. Natl Acad. Sci. USA* **108,** 9601–9606 (2011).
93. Chen, G. Y., Liu, M., Wang, F., Bertin, J. & Nunez, G. A functional role for Nlrp6 in intestinal inflammation and tumorigenesis. *J. Immunol.* **186,** 7187–7194 (2011).
94. Hu, B. *et al.* Inflammation-induced tumorigenesis in the colon is regulated by caspase-1 and NLRC4. *Proc. Natl Acad. Sci. USA* **107,** 21635–21640 (2010).
95. Hoffman, H. M. & Wanderer, A. A. Inflammasome and IL-1β-mediated disorders. *Curr. Allergy Asthma Rep.* **10,** 229–235 (2010).
96. Saleh, M. & Trinchieri, G. Innate immune mechanisms of colitis and colitis-associated colorectal cancer. *Nature Rev. Immunol.* **11,** 9–20 (2011).
97. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472,** 57–63 (2011).
98. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457,** 480–484 (2009).
99. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473,** 174–180 (2011).
100. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107,** 14691–14696 (2010).

**Acknowledgements** We thank S. Eisenbarth, M. de Zoete, C. Thaiss, P. Licona-Limón, V. Nagy and A. Rongvaux for critically reading the manuscript, and F. Manzo for help with its submission. We apologize to those whose work was not cited owing to space constraints. T.S. and J.H.M. are supported by postdoctoral fellowships from the Leukemia and Lymphoma Society. E.E. is supported by the Cancer Research Institute (2010–2012) and the United States–Israel Educational Foundation (2009), and is the recipient of the Claire and Emmanuel G. Rosenblatt award from the American Physicians for Medicine in Israel Foundation. R.F. is an Investigator of the Howard Hughes Medical Institute.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to R.F. (richard.flavell@yale.edu).

doi:10.1038/nature10760

# The DNA damage response and cancer therapy

Christopher J. Lord[1*] & Alan Ashworth[1*]

**Genomic instability is one of the most pervasive characteristics of tumour cells and is probably the combined effect of DNA damage, tumour-specific DNA repair defects, and a failure to stop or stall the cell cycle before the damaged DNA is passed on to daughter cells. Although these processes drive genomic instability and ultimately the disease process, they also provide therapeutic opportunities. A better understanding of the cellular response to DNA damage will not only inform our knowledge of cancer development but also help to refine the classification as well as the treatment of the disease.**

Towards the end of his life, the German biologist Theodor Boveri (1862–1915) published a prophetic and much-quoted manuscript on the genetic basis of cancer[1,2] that in contemporary terms[2] would be hailed as a great example of 'translation' or 'impact'. In his study of chromosome and cell behaviour in the sea urchin, Boveri suggested that the pathogenesis of cancer might be driven by a "specific and abnormal chromosome constitution"[1,2]. A century later, genomic instability is recognized as a characteristic of most solid tumours and adult-onset leukaemias. In most cancers, the instability is obvious as alterations in chromosome number and structure, a phenotype termed chromosomal instability, and as changes to the structure of DNA, such as nucleotide substitutions, insertions and deletions. When they occur in crucial 'driver' genes (of which there are probably fewer than ten per tumour) these mutations can alter cell behaviour, confer a selective advantage and drive the development of the disease. Importantly, these mutations can also influence how the tumour will respond to therapy. Alongside key driver mutations, emerging data from cancer genome sequencing suggests that a typical tumour may contain many thousands of other genetic changes. These 'passenger' mutations do not contribute directly to the disease, but are probably collateral damage from exposure to various environmental factors or defects in the molecular mechanisms that maintain the integrity of the genome.

## DNA damage and the DNA damage response

Given the potentially devastating effects of genomic instability, cells have evolved an intricate series of interlocking mechanisms that maintain genomic integrity[3]. The size of this task is daunting; the integrity of DNA is continually challenged by a variety of agents and processes that either alter the DNA sequence directly or cause mutation when DNA is suboptimally repaired. For example, the ultraviolet component of sunlight can cause up to $1 \times 10^5$ DNA lesions per cell per day[4], many of which are pyrimidine dimers. If left unrepaired, dimers that contain cytosine residues are prone to deamination, which can ultimately result in cytosine being replaced with thymine in the DNA sequence. Likewise, ionizing radiation (for example, from sunlight or cosmic radiation) can cause single-strand breaks (SSBs) and double-strand breaks (DSBs) in the DNA double helix backbone. If misrepaired — for example, the inaccurate rejoining of broken DNA ends at DSBs — these breaks can induce mutations and lead to widespread structural rearrangement of the genome. Environmental agents that cause DNA damage and mutation include cigarette smoke (40 years of heavy smoking is estimated to cause as many as 1,000 DNA lesions per cell[5]), industrial chemicals, mustard gases used in warfare and a wide variety of drugs used in chemotherapy. This final group of agents has been the mainstay of cancer therapy for decades and includes drugs that induce covalent crosslinks between DNA bases (for example, cisplatin, carboplatin, oxaliplatin and mitomycin C), drugs that attach alkyl groups to bases (such as methyl methanesulphonate and temozolomide) and agents that cause SSBs or DSBs by trapping topoisomerase I or II enzymes on DNA (such as camptothecin and etoposide)[3]. Endogenous processes also present a challenge to the integrity of the genome. For example, normal cellular metabolism generates reactive oxygen species that can oxidize DNA bases and cause SSBs. Similarly, errors in DNA replication caused by deoxyribonucleoside 5′-triphosphate (dNTP) misincorporation are potentially mutagenic, as is base depurination and deamination, an intrinsic chemical liability of DNA[3].

The variety and frequency of DNA lesions are matched by the complexity of mechanisms that counteract these threats to genomic integrity. Collectively, these mechanisms are known as the DNA damage response (DDR). In general terms, the DDR can be divided into a series of distinct, but functionally interwoven, pathways, which are defined largely by the type of DNA lesion they process (Fig. 1). Most DDR pathways encompass a similar set of tightly coordinated processes: namely the detection of DNA damage, the accumulation of DNA repair factors at the site of damage and finally the physical repair of the lesion. Most of the subtle changes to DNA, such as oxidative lesions, alkylation products and SSBs, are repaired through a series of mechanisms that is termed base excision repair (BER). In BER, damaged bases are first removed from the double helix, and the 'injured' section of the DNA backbone is then excised and replaced with newly synthesized DNA[6]. Key to this process are the enzymes poly(ADP-ribose) polymerase 1 and 2 (PARP1 and PARP2), which act as sensors and signal transducers for lesions, such as SSBs. In addition to BER, the pool of deoxynucleotides (deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxyguanosine triphosphate (dGTP) and deoxycytidine triphosphate (dCTP)) that provide the building blocks of DNA can be chemically modified before they are incorporated into the double helix. The nucleotide pool is, therefore, continually 'sanitized' by enzymes such as nudix-type motif 5 (NUDT5)[7].

Whereas small base adducts are repaired by BER, some of the bulkier single-strand lesions that distort the DNA helical structure, such as those caused by ultraviolet light, are processed by nucleotide excision repair (NER)[8]. NER is often subclassified into transcription-coupled NER, which occurs where the lesion blocks, and is detected

---

[1]The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK. *These authors contributed equally to this work.

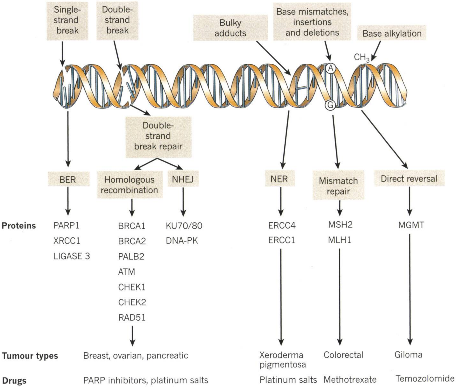


**Figure 1 | A panoply of DNA repair mechanisms maintains genomic stability.** DNA is continually exposed to a series of insults that cause a range of lesions, from single-strand breaks (SSBs) to base alkylation events. The choice of repair mechanism is largely defined by the type of lesion, but factors such as the stage of the cell cycle also have a role. Key proteins involved in each DDR mechanism, the tumour types usually characterized by DDR defects and the drugs that target these defects are shown. BER, base excision repair; NER, nucleotide excision repair; NHEJ, non-homologous end-joining. Figure modified, with permission, from ref. 72.

by elongating RNA polymerase; and global-genome NER, in which the lesion is detected not as part of a blocked transcription process but because it disrupts base pairing and distorts the DNA helix. Although these processes detect lesions using different mechanisms, they repair them in a similar way: DNA surrounding the lesion is excised and then replaced using the normal DNA replication machinery. Excision repair cross-complementing protein 1 (ERCC1) is key to this excision step.

The major mechanisms that cope with DSBs are homologous recombination[9] and non-homologous end joining (NHEJ)[10]. Homologous recombination acts mainly in the S and G2 phases of the cell cycle and is a conservative process in that it tends to restore the original DNA sequence to the site of damage. Part of the DNA sequence around the DSB is removed (known as resection) and the DNA sequence on a homologous sister chromatid is used as a template for the synthesis of new DNA at the DSB site. Crucial proteins involved in mediating homologous recombination include those encoded by the *BRCA1*, *BRCA2*, *RAD51* and *PALB2* genes. In contrast to homologous recombination, NHEJ occurs throughout the cell cycle. Rather than using a homologous DNA sequence to guide DNA repair, NHEJ mediates repair by directly ligating the ends of a DSB together. Sometimes this process can cause the deletion or mutation of DNA sequences at or around the DSB site. Therefore, compared with homologous recombination, NHEJ, although mechanistically simpler, can often be mutagenic.

Mismatch repair[11] is crucial to the DDR. It deals primarily with dNTP misincorporation and formation of 'insertion and deletion' loops that form during DNA replication. These errors cause base 'mismatches' in the DNA sequence (that is, non-Watson-Crick base pairing) that distort the helical structure of DNA and so are recognized as DNA lesions. The recognition of this distortion triggers a procession of events resulting in the excision of newly synthesized DNA encompassing the mismatch site and the resynthesis of DNA in its place. Key to the process of mismatch repair are proteins encoded by the mutS and mutL homologue genes, such as *MSH2* and *MLH1*.

Finally, translesion synthesis and template switching allow DNA to continue to replicate in the presence of DNA lesions that would otherwise halt the process. Translesion synthesis and template switching are therefore usually considered to be part of the DDR. In translesion synthesis, relatively high-fidelity DNA replication polymerases are transiently replaced with low-fidelity 'translesion' polymerases that are able to synthesize DNA using a template strand encompassing a DNA lesion. Once the replication fork passes the site of the lesion, the low-fidelity DNA polymerases are normally replaced with the usual high-fidelity enzyme, which allows DNA synthesis to continue as normal. In template switching, the DNA lesion is bypassed at the replication fork by simply leaving a gap in DNA synthesis opposite the lesion. After the lesion has passed the replication fork, the single-strand gap is repaired using template DNA on a sister chromatid, similar to the process used during homologous recombination.

Although sometimes considered distinct from the DDR, the mechanisms that control the integrity of telomeric DNA at the end of each human chromosome also act as a barrier against genomic instability and mutation[12]. Rather than being an exposed DNA double-helix structure at the end of the chromosome, telomeric DNA comprises a series of guanine-rich, repetitive DNA sequences. These enable the telomeric DNA to be bound in a loop-like structure with a series of proteins (telomere repeat binding factor 1 and 2 (TERF1 and TERF2), protection of telomeres 1 (POT1), TERF1-interacting nuclear factor 2 (TINF2)) that form a shelterin complex. This 'capping' structure prevents the otherwise exposed ends of different chromosomes from becoming fused together (a process known as end–end fusion) by the DDR. In most somatic cells, the length of telomeric DNA is reduced at each cell cycle, a process termed telomere attrition. Eventually, telomeres reach a crucial length that precludes the formation of an effective shelterin complex. In normal cells, this failure in telomere capping induces a p53-mediated response that results in cellular senescence, a mechanism that ultimately prevents unlimited cell proliferation. When the p53 response is abrogated, end–end fusions occur, which leads to the formation of chromosomes with two centromeres. At mitosis, each centromere in a dicentric chromosome attaches to an opposite spindle pole. The physical stress of opposing forces during chromosome segregation shears dicentric chromosomes, resulting in broken chromosome ends. This process is an ideal substrate for the chromosome translocation, focal DNA amplification and deletion events that potentially drive tumorigenesis. Sheared chromosomes could also be the source of the mutagenic 'big bang' or, chromothripsis[13], that has been posited to occur in some tumours. Whether or not this is the case, compared with healthy cells, many pre-invasive and invasive cancers have significantly shorter telomeres, a situation in which end–end fusions and dicentric chromosome shearing could easily occur[12].

The core DDR machinery does not work alone, but is coordinated with a set of complementary mechanisms that are also crucial to maintaining the integrity of the genome. For example, chromatin-remodelling proteins allow the DNA repair apparatus to gain access to the damaged DNA[14]. Similarly, processes such as homologous recombination are dependent on the close proximity of a sister chromatid, and probably work in tandem with the chromosome cohesion machinery to ensure this is the case. Crucially, DDR core components interact with the cell-cycle-checkpoint and chromosome-segregation machinery. These interactions allow DNA repair to occur before mitosis takes place and ensure that the correct complement of genetic material is passed on to daughter cells[15].

## Cancer

One of the most pervasive characteristics of human tumours is genomic instability[16]. Although the specific DDR defects are not known in most cancers, there are several examples in which there is an incontrovertible link between a particular DDR dysfunction and the neoplastic phenotype. For example, 15% of sporadic colorectal tumours show an abnormal shortening or lengthening of dinucleotide repeat sequences. This DNA mutation pattern, known as microsatellite instability, is probably caused by an inability to repair DNA replication errors when mismatch repair is defective. Microsatellite instability is observed not only in sporadic colorectal tumours but also in a familial form of the disease known as hereditary non-polyposis colorectal cancer (HNPCC). HNPCC is associated with loss-of-function mutations in mismatch repair genes, such as *MSH2* and *MLH1* (ref. 11).

Similarly, a subset of both sporadic and familial cancers seems to be associated with defects in homologous recombination. Integrated (genomic, transcriptomic, methylomic) analysis of just under 500 samples of high-grade serous ovarian adenocarcinoma has estimated that homologous recombination is probably defective in approximately half. These homologous recombination defects are largely driven by DNA mutations or epigenetic silencing of genes such as *BRCA1* and *BRCA2* (ref. 17). Familial forms of breast, ovarian and pancreatic cancer are also associated with loss-of-function mutations in homologous recombination modifying genes, such as *BRCA1*, *BRCA2*, *PALB2*, *ATM*, *RAD51C* and *RAD51D*. A number of rare, syndromic conditions also exist in which the underlying genetic determinant is a loss-of-function mutation in a gene that modulates the response to DSBs[9]. These conditions include ataxia telangiectasia (ATM), Bloom's syndrome, Fanconi anaemia, Rothmund–Thomson syndrome, Werner syndrome and Nijmegen breakage syndrome, all of which have variable clinical phenotypes but are similar in that they cause extreme radiosensitivity, a characteristic of the inability to process DSBs effectively.

Evidence is also building that the DDR is not only invoked but also dysfunctional at an early stage in the development of neoplasia. Markers of DSBs, such as nuclear γH2AX foci (a histone phosphorylation event that occurs on chromatin surrounding a DSB), are markedly elevated in some precancerous lesions[18,19]. One hypothesis[19] proposes that the original cause of these effects is oncogene activation. The activation of oncogenes such as *MYC* and *RAS* stimulates the firing of multiple replication forks as part of a proliferative program. These forks rapidly stall, collapse and form DSBs because they exhaust the available dNTP pool or because multiple forks collide on the same chromosome. Regardless of the mechanism, stalled and collapsed forks normally invoke the DDR and cell-cycle checkpoints that enable DNA lesions to be repaired before mitosis takes place. For precancerous lesions to progress to mature tumours, it is thought that critical DSB signal transduction and cell-cycle checkpoint proteins, such as ATM, ataxia telangiectasia and Rad3 related (ATR), and the master 'gatekeeper' protein p53 become inactivated. With these DDR components rendered dysfunctional, collapsed forks are not effectively repaired, and cells proceed through the cell cycle with DNA lesions intact, increasing the chance of mutagenesis[18,19].

## Selection of cancer treatment

Historically, DDR defects have already been exploited in the form of many commonly used chemotherapy and radiotherapy regimens. One example is the use of platinum salts (carboplatin or cisplatin) frequently given in combination with the taxane paclitaxel in patients with advanced ovarian cancer. Platinum salts cause DNA inter- and intrastrand crosslinks — lesions that are recognized by the DDR and repaired by a combination of NER and homologous recombination. These agents may be effective in patients with ovarian cancer because these tumours commonly harbour defects in homologous recombination[20], especially high-grade serous ovarian cancers. Estimates vary, but somewhere between 37% and 55% of high-grade serous ovarian cancers have germ line or somatic mutations in *BRCA1* or *BRCA2* and some have epigenetic inactivation of *BRCA1* (ref. 21). Cisplatin also improves survival in patients with ERCC1-negative non-small-cell lung cancers, but not in those with ERCC1-positive tumours[22]. The NER pathway, and in particular the ERCC1–ERCC4 complex, is essential for the removal of platinum adducts, and low levels of ERCC1 are associated with an accumulation of adducts and an increased apoptotic response to cisplatin. Finally, where alkylating agents work, they probably also exploit a DDR defect. Normally, O6-methyltransferase (MGMT) directly reverses DNA alkylation events caused by drugs such as temozolomide when they occur on the O6 position of guanine residues. However, in some tumours, the promoter of the *MGMT* gene is methylated, causing a reduction in MGMT expression. This defect causes $O^6$ alkylation events formed by temozolomide to persist so that they eventually pair incorrectly with thymidine residues during replication. The resultant DNA mismatch is recognized by the mismatch repair apparatus, which ultimately stimulates cell death. *MGMT* methylation has also been implicated in the response of tumours (such as gliomas) to temozolomide, and is now being used as a biomarker to direct the use of therapy[23] (Table 1).

An alternative approach to chemotherapy is the design of drugs that target specific DDR components. Topoisomerase inhibitors, such as irinotecan (a topoisomerase I inhibitor) and etoposide (a topoisomerase II inhibitor), could be regarded as the first generation of DDR 'targeted' agents[24]. Topoisomerases catalyse the breaking and rejoining of the phosphodiester backbone as part of the process that unwinds the torsional structure of DNA before processes such as transcription and DNA replication; agents that inhibit this function, such as irinotecan and etoposide, leave DNA breaks across the genome and have been used effectively in some cancers (Table 1). The development of the next generation of DDR inhibitors, particularly PARP inhibitors, has given considerable impetus to this area.

### PARP inhibitors as targeted therapy

PARP1 and PARP2 are members of the PARP protein superfamily. These enzymes catalyse the polymerization of ADP-ribose moieties onto target proteins (PARsylation) using $NAD^+$ as a substrate, releasing nicotinamide in the process. This modification often modulates the conformation, stability or activity of the target protein[25]. The best understood role of PARP1 is in SSB repair ((SSBR) a form of BER). PARP1 initiates this process by detecting and binding SSBs through a zinc finger in the PARP protein. Catalytic activity of PARP1 results in the PARsylation of PARP1 itself and the PARsylation of a series of additional proteins, such as XRCC1 and the histones H1 and H2B; when PARP activity is inhibited, SSB repair is compromised[25]. The development of PARP inhibitors began with the observation that nicotinamide, a product of PARP catalytic activity, is itself a weak PARP inhibitor, as are nicotinamide analogues, such as 3-aminobenzamide. The screening of chemical libraries and the subsequent chemical refinement

**Table 1 | Examples of DDR inhibitors in clinical use or in development**

| Class | Name | Stage of development | Cancer type |
|---|---|---|---|
| Topoisomerase I inhibitor | Irinotecan | Licensed for use | Mainly colorectal cancers |
| Topoisomerase II inhibitor | Etoposide phosphate | Licensed for use | For example, testicular cancer, Ewing's sarcoma, lung cancer, lymphoma, lymphocytic leukaemia and glioblastoma multiforme |
| DNA protein kinase inhibitor | CC-115 | Phase I study under way | Advanced solid tumours, non-Hodgkin's lymphoma and multiple myeloma |
| PARP inhibitor | Olaparib | 2 phase II studies completed, 9 additional phase II studies under way | Ovarian, breast, gastric, colorectal and pancreatic tumours and a range of other solid tumours |
| | Veliparib | 16 phase II studies under way | Ovarian, breast, gastric, colorectal and pancreatic tumours and a range of other solid tumours |
| | Rucaparib | 2 phase II studies recruiting | Breast and other solid tumours |
| | MK4827 | Phase I study completed, phase Ib ongoing | Glioblastoma multiforme, melanoma and other solid tumours and leukaemias |
| CHEK1 inhibitor | AZD7762 | Phase I study completed | Range of solid tumours |
| MGMT inhibitor | Lomeguatrib | Phase I study under way | Colorectal cancer and other solid tumours |
| Telomerase inhibitor | Imetelstat | 4 phase II studies under way | Solid tumours and leukaemias |

of compounds have led to second generation PARP inhibitors (for example, the benzimidazole-4-carboxamide NU1025), and to a third generation of clinically usable inhibitors, such as the tricyclic lactam indole, AG014699, and phthalazinones, such as AZD2281 (olaparib)[26].

For cancer therapy, PARP inhibitors are potent chemosensitizers. For example, 3-aminobenzamide enhances the cytotoxic effect of the DNA-methylating agent dimethyl sulphate, and several preclinical studies have shown that clinical PARP inhibitors increase the effects of another DNA-methylating agent, temozolomide[25,26]. In 2005, potent PARP inhibitors were shown to selectively inhibit the growth of cells with defects in either the *BRCA1* or *BRCA2* genes, suggesting a new use for these agents[27,28]. PARP inhibition probably works by allowing the persistence of spontaneously occurring SSBs, or by inhibiting PARP release from a DNA lesion. Whichever is the case, both of these DNA lesion types could conceivably stall and collapse replication forks, potentially creating lethal DSBs[29]. In normal cells, the effects of PARP inhibition are buffered by homologous recombination, which repairs the resultant DSB. However, effective homologous recombination is reliant on functioning *BRCA1* and *BRCA2*, so when these genes are defective — as they are in tumours of germline *BRCA*-mutant carriers — DSBs are left unrepaired, and potent PARP inhibitors can cause cell death. In some *in vitro* models, cells with *BRCA* mutations are more than 1,000 times more sensitive to PARP inhibitors than cells with wild-type alleles[27]. This genotype-specific selectivity provides a strong argument for these agents to be tested clinically as single agents, rather than in combination with chemotherapies. Moreover, the sensitivity to PARP inhibitors seems to be defined more by the *BRCA* genotype of a cancer cell than by its tissue of origin. Breast, ovarian and prostate cancers with *BRCA* mutations all seem to be profoundly sensitive to these drugs. The demonstration that *BRCA*-mutant cells are sensitive to PARP inhibitors has now become the emblematic 'poster child' for exploiting synthetic lethality in cancer therapy. Synthetic lethality, a genetic concept developed by Dobzhansky[30], Lucchesi[31] and others and championed as a therapeutic approach for cancer by Hartwell and colleagues[32] and Kaelin[33], is the process in which different single gene defects are compatible with cell viability, but the synthesis or combination of these gene defects results in cell death. In this case, *BRCA* dysfunction is synthetically lethal when PARP is inhibited. Furthermore, the demonstration that one DDR pathway (homologous recombination) could be targeted by inhibition of another (PARP-mediated BER and SSB repair) suggests that DDR pathways functionally buffer one another. This hypothesis has long been established in model organisms (such as yeast) in which synthetic lethalities can be easily identified using genetic screens. These studies and the BRCA/PARP paradigm in human tumour cells has inspired the search for other DDR synthetic lethalities that could be exploited in the treatment of cancer, including other synthetic lethal approaches for *BRCA*-mutant tumour cells[34], as well as for tumour cells with other gene defects. For example, the mismatch repair gene defects found in colorectal cancer are synthetically lethal with inhibition of genes controlling the response to oxidative DNA damage, such as *POLB*[35], *POLG*[35] and *PINK1* (ref. 36). The first therapeutic test of the hypothesis that mismatch-repair defective cancers can be treated with a synthetic lethality approach[37] is now in a proof-of-concept clinical trial (http://www.clinicaltrials.gov, NCT00952016).

The preclinical work described above has paved the way for subsequent clinical studies of PARP inhibitors (Table 1). Phase I studies have established the safety of olaparib, a potent PARP inhibitor, as a single agent and shown that significant and durable antitumour responses can be established in patients with *BRCA*-mutant breast, ovarian or prostate tumours[38]. Furthermore, olaparib does not seem to cause many of the side effects associated with standard chemotherapies. In subsequent phase II clinical studies, 40% of patients with breast or ovarian cancer with germline BRCA mutations had a favourable response to the drug. This is a particularly high response given that the patients in these trials had been heavily pretreated and had become resistant to a range of chemotherapies. Such work has led to a series of additional PARP inhibitors being tested in clinical trials[39].

Recently, attention has focused on clinical trials assessing iniparib. Unlike drugs such as olaparib and AG014699, iniparib was originally proposed to be a non-competitive inhibitor of PARP1 that disrupts the interaction between PARP1 and DNA[40]. In a phase II trial, iniparib elicited a significant response in patients with triple negative (oestrogen-receptor, progesterone-receptor and *HER2*-amplification negative) breast cancer when used in combination with gemcitabine and carboplatin. However, a subsequent phase III trial failed to deliver the predefined improvements in progression-free or overall survival. The reasons for this failure are unclear, although differences in the patient populations and study design may have contributed. However, the clinical activity of iniparib may not be related to PARP inhibition[41]; this drug has very low PARP inhibition *in vitro,* and its mechanism of action *in vivo* is unclear[40,41]. We therefore believe that there is little evidence to regard the relative failure of iniparib in phase III trials as a PARP-inhibitor class effect. In fact, the clinical observations with potent catalytic-site PARP inhibitors have been promising when used in appropriate populations[39].

Alongside the clinical trials of PARP inhibitors, preclinical research has extended our understanding of how these agents work and how they might be used most effectively in the clinic. Since the initial reports of

BRCA/PARP synthetic lethality, several other determinants of PARP-inhibitor sensitivity have been identified. As expected, most of these determinants cause PARP-inhibitor sensitivity by modulating homologous recombination (for example, RAD51, ATRX, SHFM1, RPA1, NBN, ATR, ATM, CHEK1, CHEK2, miR-182, CDK1, SWI5–SFR1, USP1/UAF1 and several Fanconi anaemia proteins[42–47]). Of particular interest is the observation that cyclin dependent kinase 1 (CDK1) inhibition suppresses homologous recombination and may therefore elicit PARP-inhibitor sensitivity both *in vitro* and *in vivo*[46]. More extensive studies are now required to identify the patient subgroup, and specifically the genetic context, in which this combination could be used. In addition, mutations in the tumour-suppressor gene *PTEN*[48] and the *TMPRSS2–ERG* fusion[49] can sensitize cells to clinical PARP inhibitors; both of these changes are relatively common in human tumours so could be used as biomarkers to direct the use of these agents. In the case of the PTEN/PARP synthetic lethality, one model suggests that PTEN deficiency causes a downregulation in RAD51 levels, therefore disrupting homologous recombination[49]. Finally, several PARP-inhibitor-resistance-causing mechanisms have been identified, particularly in *BRCA1*- or *BRCA2*-mutant tumour models. In two studies, PARP-inhibitor resistance was caused by secondary mutations in the *BRCA1* or *BRCA2* genes themselves[50,51]. These mutations restored the open reading frame of the *BRCA1* or *BRCA2* gene, enabled the generation of functional proteins and restored the ability to repair DNA damage caused by PARP inhibitors. Importantly, these secondary mutations were observed not only in *in vitro* models but also in ovarian tumours from patients who had become resistant to platinum salts such as carboplatin[50–53]. These drugs target cells with deficient homologous recombination, albeit with less efficiency than PARP inhibitors. Subsequent studies have suggested that these reversion events are relatively common in *BRCA*-mutant cancers that do not respond to platinum salts. This is especially true in ovarian cancer, in which more than 42% of the carboplatin-resistant tumours tested had secondary mutations that restored the *BRCA1* or *BRCA2* open reading frames[54]. Taken together, these observations could help to explain why some platinum-refractory *BRCA*-mutant patients in PARP-inhibitor clinical trials failed to respond and strengthen the argument for the inclusion of biomarkers that monitor *BRCA1* and *BRCA2* sequence and homologous recombination function in future trials. Furthermore, the restoration of homologous recombination and drug resistance by secondary mutations provides strong confirmation that, at least in this setting, PARP inhibitors and platinum salts require defective homologous recombination for antitumour activity. Other identified mechanisms of PARP-inhibitor response include the upregulation of genes that encode P-glycoprotein efflux pumps[55] and the loss of TP53BP1 (a DDR protein) in *BRCA1*-deficient cells, suggesting that these mechanisms should also be assessed in clinical biopsies[56,57].

In addition to dissecting the mechanisms behind PARP-inhibitor sensitivity and resistance, there has been considerable debate regarding the potential of PARP inhibitors in non-familial cancers. Germline mutations in *BRCA1* or *BRCA2* are common in hereditary breast cancer, but relatively rare in sporadic breast cancer. However, many sporadic tumours have molecular and histopathological features in common with BRCA-deficient tumours, a concept originally termed BRCAness[20]. For example, BRCAness may be caused by methylation of the *BRCA1* gene promoter (which causes epigenetic silencing of BRCA1 expression) or even somatic mutation of genes that modulate homologous recombination. Furthermore, traditional histopathological methods and gene-expression profiling approaches have suggested some level of phenotypic or molecular overlap between triple-negative breast cancers, basal-like breast cancers and *BRCA1*-mutant familial breast cancers[20]. Although the overlap is not absolute, it has led to the hypothesis that there may be a subset of sporadic breast cancers that exhibit features of BRCAness and may respond favourably to PARP inhibitors[20]. This hypothesis has yet to be fully tested experimentally, in part because existing tumour-cell models of triple-negative breast cancer do not encompass the molecular and functional heterogeneity of triple-negative breast cancer. In the clinical setting, patients with triple-negative breast cancer do not show a frequent response to either olaparib or veliparib[39], although the number of patients tested is still small. Needless to say, larger studies are needed, as are clinical trials that lead to a full drug licence for PARP inhibitors in *BRCA*-mutation-related breast cancer.

## Future issues

More research is needed into how DDR activity affects the clinical efficacy of DNA-damaging chemotherapies, and here we highlight some areas that need to be addressed in the future.

### Pharmacological targeting

Although the DDR has been the subject of intense study, the biochemical and genetic interactions between the various DNA repair pathways is not completely understood. This understanding would be facilitated by the availability of potent and specific DDR-pathway small-molecule inhibitors. Past efforts have focused on the protein kinases that transduce DDR signals (such as ATM, ATR and DNA-PK) and on the PARP superfamily members that detect damage and act as signal transduction mediators. For example, synthetic lethalities with ATM and DNA-PK suggest that ATM inhibitors could increase the potential of DNA-damaging chemotherapy in *TP53*-mutant tumours, and DNA-PK inhibitors could be used in combination with chemotherapy in *ATM*-mutant tumours[58]. A number of other enzyme classes have also been identified as having a role in the DDR, including enzymes that mediate SUMOylation, histone modification, acetylation and ubiquitylation[59]. The availability of pharmacological inhibitors of these proteins would facilitate the search for additional synthetic lethal interactions. An example is the identification of small-molecule inhibitors (such as MLN4924) that modulate neddylation, a post-translational modification process analogous to ubiquitylation[60]. Inhibition of neddylation, which is induced by MLN4924, causes stabilization of the DNA-replication licensing factor CDT1, DNA synthesis in the absence of cell-cycle progression and activation of the DDR[60]. These agents could potentially target tumour cells that have existing DNA repair defects or are under replicative stress.

The clinical use of DDR-pathway drugs remains a challenge. Addition of such agents to standard chemotherapy regimens will probably lead to greater side effects, and therefore more work in this area is needed on the appropriate dosing and scheduling of inhibitors. These studies should also address the possibility that targeting the DDR could also increase the rate of second cancers or leukaemias, something that has already been observed for some chemotherapies. In principle, these iatrogenic effects might be avoided if DDR inhibitors could be administered for a relatively short duration, a hypothesis that should be tested clinically as well as in preclinical models. Finally, combining DDR inhibitors with other targeted agents, such as signal transduction inhibitors or targeted radiation, has been underexplored and could form the basis of effective treatment regimens.

### Cancer stem cells and the tumour micro-environment

Many human cancers contain cells that share some of the self-renewal and differentiation properties of stem cells. These cancer stem cells are thought to act as tumour-initiating cells that, after treatment, may potentially repopulate the therapy-resistant tumour. With this role in mind, understanding the response to DNA-damaging therapy in cancer stem cells may prove essential in understanding whether a treatment will ultimately work. Evidence is accumulating that the DDR has distinct properties in cancer stem cells. These may include controversial and unproven concepts such as the 'immortal strand' model in which parental, but not newly synthesized DNA, is retained in stem cells after mitosis; retaining only the parental strand of DNA is thought to limit the impact of DNA replication errors to the stem cell population[61]. Whether this is the case or not, a number of studies have suggested that resistance to therapy is mediated by altered DDR activity in cancer stem cells and tumour-initiating cells. For example, a comparison of

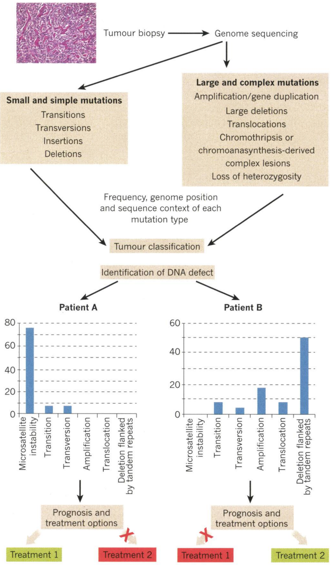

Tumour biopsy
Genome sequencing
Small and simple mutations
Transitions
Transversions
Insertions
Deletions
Large and complex mutations
Amplification/gene duplication
Large deletions
Translocations
Chromothripsis or chromoanasynthesis-derived complex lesions
Loss of heterozygosity
Frequency, genome position and sequence context of each mutation type
Tumour classification
Identification of DNA defect
Patient A
80
60
40
20
0
Microsatellite instability
Transition
Transversion
Amplification
Translocation
Deletion flanked by tandem repeats
Patient B
60
40
20
0
Microsatellite instability
Transition
Transversion
Amplification
Translocation
Deletion flanked by tandem repeats
Prognosis and treatment options
Treatment 1
Treatment 2
Prognosis and treatment options
Treatment 1
Treatment 2

**Figure 2 | Genomic scars in cancer.** Modern sequencing approaches mean that the nucleotide sequence and mutational spectrum of tumour biopsies can be rapidly determined. This spectrum, or sequence scar, could be used to classify each tumour. Classifiers could take into account the frequency and location of mutations, as well as their predominant sequence context (for example, insertions and deletions that are flanked by small regions of microhomology[13,73]). Such classifiers could then be used to estimate the probable outcome for the patient and to select the most appropriate therapy. The choice of therapy is defined by the balance of sequence scars found in tumour genome biopsies.

breast tumour biopsies before and after chemotherapy treatment has shown an enrichment of cancer stem cells, which suggests an intrinsic resistance of these cells. In mouse models, cancer stem cells deficient in p53 showed accelerated DNA repair activity[62], as well as high levels of AKT and WNT signalling, both of which promoted cancer-stem-cell survival after exposure to ionizing radiation[63]. In glioblastoma multiforme, cancer stem cells have been characterized on the basis of expression of prominin (CD133; now known as PROM1)[64]. CD133$^+$ cells are enriched after irradiation of human glioblastoma multiforme xenografts, suggesting that cancer stem cells are necessary for tumour regrowth after radiotherapy. Cancer stem cells showed more robust activation of the DDR and higher intrinsic DNA repair activity. Blocking the kinases CHEK1 and CHEK2 led to enhanced killing of cancer stem cells by ionizing radiation[65], suggesting that a pharmacological approach may overcome this intrinsic resistance. Much more detailed and systematic knowledge of the consequences of DNA damage and the DDR in stem cells is required to exploit this important area.

If we are to optimize therapies that exploit the DDR, then the impact of the tumour microenvironment and the non-tumour cells that sit within a neoplastic lesion must be considered. For example, non-tumour stromal cells form at least half of the cellular complement of pancreatic ductal adenocarcinomas. This stromal compartment contains pancreatic stellate cells, which, in *in vitro* systems at least, cause tumour cells to become resistant to ionizing radiation; this radioprotective effect could explain the relative radioresistance of pancreatic ductal adenocarcinomas. Already, there is a preliminary understanding of how stromal and tumour cells interact and communicate to protect against ionizing radiation[66], but the understanding of how the microenvironment modulates the DDR is still limited. Given the complexity of studying molecular mechanisms *in vivo* and of effectively modelling cell–cell interactions *in vitro*, this is not surprising. These technical problems are not insurmountable, but they do raise an important issue: historically, much of our knowledge of the DDR has originated from studies of cells *in vitro*. Understanding the DDR *in vivo* is critical if cancer therapies are to be modelled correctly.

## The genomic scar

Historically, cancer taxonomy has mainly been based on the anatomical site of a tumour and its clinical and histopathological features. Cancers are now increasingly classified according to their underlying pattern of pathogenic mutations, a process which is being facilitated by exome or whole-genome sequencing of each cancer[67]. Treatment decisions can then be made on the basis of pharmacologically targetable drivers of the disease. In addition to identifying the driver genes affected, examining the collateral genomic damage caused by DNA repair deficiency may help in the classification of cancer and could provide biomarkers to inform therapy. In essence, the mutational spectrum of a tumour, revealed by its DNA sequence, is a sequence 'scar' that reflects both the mutagens that a tumour has been exposed to and the repair processes that have operated to mitigate their impact[67]. Already, there are some precedents to suggest that this approach could have merit. For example, homologous-recombination deficient tumours have a characteristic mutational spectrum defined by the use of error-prone DSB repair mechanisms[68,69]. Therefore, the wider tumour genome sequence or sequence scar in a tumour with defective homologous recombination might be used to decide which patients should be treated with PARP inhibitors. Microsatellite instability is another example. In many cases, microsatellite instability in a tumour reveals both the mutagenic nature of DNA replication and the failure to reverse these sequence errors when mismatch repair is defective[11]. Given the impact that the DDR can have on both disease development and response to treatment, it seems reasonable to consider categorizing tumours according to their DNA repair defects and to use this information to personalize treatment (Fig. 2). As a result of the speed at which in-depth, whole-genome tumour sequences are now being generated, we will soon have a first draft of the mutational landscape of all human cancers, which will allow these possibilities to be addressed.

To match these molecular signatures of DNA mutation and repair, we require functional biomarkers that identify DDR defects and aid in the selection of therapy. Again, there are preliminary signs that this could be achievable: a mechanistic determinant of PARP-inhibitor sensitivity, RAD51 foci deficiency, can be measured by immunohistochemical staining in tumour biopsies[70], and the ability to detect other DNA repair complexes by immunohistochemistry is improving[71]. In principle, functional biomarkers should be identifiable for most DDR pathways even though significant technical and practical issues need to be resolved before such tests can be used in the clinic. Because the DDR pathway is fully engaged only after DNA is damaged, standard tumour biopsies or specimens are not particularly useful for measurement of DDR-pathway activity. A functional test of DDR-pathway deficiency might therefore need to include some form of genotoxic stress, followed by a measurement of the molecular response. For example, to test the homologous recombination defect in tumour cells by measuring the RAD51 response, cells must first be exposed to an agent that causes DNA damage. These agents could include ionizing radiation, chemotherapy or the PARP inhibitor itself. Clinically, this could mean that patients need to be treated with a DNA-damaging agent before biopsy for the purposes of biomarker assessment or alternatively tumour biopsies would have to be exposed to DNA-damaging agents *ex vivo*. Both scenarios have practical or logistical issues. For instance, tumour cells may not respond in the same way *ex vivo* as they do *in vivo*, and pretreating a patient to assess tumour response might not always be feasible from a clinical perspective. Nevertheless, molecular and functional assays could mean that a tumour could be defined as, for example, 'deficient for homologous recombination and potentially sensitive to PARP inhibitor' before treatment rather than simply being classified as a 'non-familial, triple-negative breast cancer'. The development of functional biomarkers of other DDR pathways would considerably advance this area.

The genomic instability of tumour cells is one of their most pervasive characteristics and presents a therapeutic opportunity. Already, DNA-damaging chemotherapies have provided the core of cancer treatment for the past half-century and, if used appropriately, targeted therapies that exploit the DDR could deliver better therapeutic responses in the future. Nevertheless, there are still significant unexplored territories for the field; an understanding of the DDR *in vivo*, the definition of clinical biomarkers that allow the selection of appropriate therapy and an understanding of how the DDR interacts with the other homeostatic systems in the cell are relative unknowns. We believe that the crucial factors that have previously allowed this field to progress — the comprehensive study of the basic biology of the DDR combined with a desire to apply this information in the clinic — will drive significant advances in the future. ■

1. Boveri, T. *Zur Frage der Entstehung Maligner Tumoren* (Gustav Fischer, 1914).
2. Boveri, T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J. Cell Sci.* **121,** 1–84 (2008). **A translation of Boveri's original work that proposed the hypothesis that a form of genomic dysregulation characterizes tumour cells.**
3. Ciccia, A. & Elledge, S. J. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40,** 179–204 (2010).
4. Hoeijmakers, J. H. DNA damage, aging, and cancer. *N. Engl. J. Med.* **361,** 1475–1485 (2009).
5. Phillips, D. H., Hewer, A., Martin, C. N., Garner, R. C. & King, M. M. Correlation of DNA adduct levels in human lung with cigarette smoking. *Nature* **336,** 790–792 (1988).
6. David, S. S., O'Shea, V. L. & Kundu, S. Base-excision repair of oxidative DNA damage. *Nature* **447,** 941–950 (2007).
7. Nakabeppu, Y. *et al.* Mutagenesis and carcinogenesis caused by the oxidation of nucleic acids. *Biol. Chem.* **387,** 373–379 (2006).
8. Cleaver, J. E., Lam, E. T. & Revet, I. Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nature Rev. Genet.* **10,** 756–768 (2009).
9. Moynahan, M. E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nature Rev. Mol. Cell Biol.* **11,** 196–207 (2010).
10. Lieber, M. R. NHEJ and its backup pathways in chromosomal translocations. *Nature Struct. Mol. Biol.* **17,** 393–395 (2010).
11. Jiricny, J. The multifaceted mismatch-repair system. *Nature Rev. Mol. Cell Biol* **7,** 335–346 (2006).
12. Artandi, S. E. & DePinho, R. A. Telomeres and telomerase in cancer. *Carcinogenesis* **31,** 9–18 (2010).
13. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144,** 27–40 (2011).
14. Bell, O., Tiwari, V. K., Thoma, N. H. & Schubeler, D. Determinants and dynamics of genome accessibility. *Nature Rev. Genet.* **12,** 554–564 (2011).
15. Warmerdam, D. O. & Kanaar, R. Dealing with DNA damage: relationships between checkpoint and repair pathways. *Mutat. Res.* **704,** 2–11 (2010).
16. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
17. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474,** 609–615 (2011).
18. Bartkova, J. *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444,** 633–637 (2006).
19. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319,** 1352–1355 (2008). **References 18 and 19 describes how oncogenes are able to induce the DDR early in oncogenesis.**
20. Turner, N., Tutt, A. & Ashworth, A. Hallmarks of 'BRCAness' in sporadic cancers. *Nature Rev. Cancer* **4,** 814–819 (2004).
21. Banerjee, S., Kaye, S. B. & Ashworth, A. Making the best of PARP inhibitors inovarian cancer. *Nature Rev. Clin. Oncol.* **7,** 508–519 (2010).
22. Friboulet, L. *et al.* Molecular characteristics of ERCC1-negative versus ERCC1-positive tumors in resected NSCLC. *Clin. Cancer Res.* **17,** 5562–5572 (2011).
23. Weller, M. *et al. MGMT* promoter methylation in malignant gliomas: ready for personalized medicine? *Nature Rev. Neurol.* **6,** 39–51 (2010).
24. Pommier, Y., Leo, E., Zhang, H. & Marchand, C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem. Biol.* **17,** 421–433 (2010).
25. Rouleau, M., Patel, A., Hendzel, M. J., Kaufmann, S. H. & Poirier, G. G. PARP inhibition: PARP1 and beyond. *Nature Rev. Cancer* **10,** 293–301 (2010).
26. Zaremba, T. & Curtin, N. J. PARP inhibitor development for systemic cancer targeting. *Anticancer Agents Med. Chem.* **7,** 515–523 (2007).
27. Farmer, H. *et al.* Targeting the DNA repair defect in *BRCA* mutant cells as a therapeutic strategy. *Nature* **434,** 917–921 (2005).
28. Bryant, H. E. *et al.* Specific killing of *BRCA2*-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434,** 913–917 (2005). **References 27 and 28 describe the synthetic lethal interaction between *BRCA* mutations and PARP inhibition.**
29. Lord, C. J. & Ashworth, A. Targeted therapy for cancer using PARP inhibitors. *Curr. Opin. Pharmacol.* **8,** 363–369 (2008).
30. Dobzhansky, T. Genetics of natural populations. Xiii. Recombination and variability in populations of *drosophila pseudoobscura*. *Genetics* **31,** 269–290 (1946).
31. Lucchesi, J. C. Synthetic lethality and semi-lethality among functionally related mutants of *Drosophila melanogaster*. *Genetics* **59,** 37–44 (1968).
32. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating genetic approaches into the discovery of anticancer drugs. *Science* **278,** 1064–1068 (1997).
33. Kaelin, W. G. Jr. Synthetic lethality: a framework for the development of wiser cancer therapeutics. *Genome Med.* **1,** 99 (2009).
34. Issaeva, N. *et al.* 6-Thioguanine selectively kills BRCA2-defective tumors and overcomes PARP inhibitor resistance. *Cancer Res.* **70,** 6268–6276 (2010).
35. Martin, S. A. *et al.* DNA polymerases as potential therapeutic targets for cancers deficient in the DNA mismatch repair proteins MSH2 or MLH1. *Cancer Cell* **17,** 235–248 (2010).
36. Martin, S. A., Hewish, M., Sims, D., Lord, C. J. & Ashworth, A. Parallel high-throughput RNA interference screens identify PINK1 as a potential therapeutic target for the treatment of DNA mismatch repair-deficient cancers. *Cancer Res.* **71,** 1836–1848 (2011).
37. Martin, S. A. *et al.* Methotrexate induces oxidative DNA damage and is selectively lethal to tumour cells with defects in the DNA mismatch repair gene *MSH2*. *EMBO Mol. Med.* **1,** 323–337 (2009).
38. Fong, P. C. *et al.* Inhibition of poly(ADP-ribose) polymerase in tumors from *BRCA* mutation carriers. *N. Engl. J. Med.* **361,** 123–134 (2009). **This article describes the results of the first clinical trial to demonstrate the clinical potential of a synthetic lethal approach for therapy.**
39. Balmaña, J., Domchek, S. M., Tutt, A. & Garber, J. E. Stumbling blocks on the path to personalized medicine in breast cancer: the case of PARP inhibitors for BRCA1/2-associated cancers. *Cancer Discov.* **1,** 29–34 (2011).
40. Mendeleyev, J., Kirsten, E., Hakam, A., Buki, K. G. & Kun, E. Potential chemotherapeutic activity of 4-iodo-3-nitrobenzamide. Metabolic reduction to the 3-nitroso derivative and induction of cell death in tumor cells in culture. *Biochem. Pharmacol.* **50,** 705–714 (1995).
41. Liu, X. et al. Iniparib non-selectively modifies cysteine-containing proteins in tumor cells and is not a bona fide PARP inhibitor. *Clin. Cancer Res.* advance ahead of print http://dx.doi.org/10.1158/1078-0432.CCR-11-1973 (29 November 2011).
42. McCabe, N. *et al.* Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res.* **66,** 8109–8115 (2006).
43. Peasland, A. *et al.* Identification and evaluation of a potent novel ATR inhibitor, NU6027, in breast and ovarian cancer cell lines. *Br. J. Cancer* **105,** 372–381 (2011).
44. Murai, J. *et al.* The USP1/UAF1 complex promotes double-strand break repair through homologous recombination. *Mol. Cell. Biol.* **31,** 2462–2469 (2011).
45. Moskwa, P. *et al.* miR-182-mediated downregulation of *BRCA1* impacts DNA repair and sensitivity to PARP inhibitors. *Mol. Cell* **41,** 210–220 (2011).
46. Johnson, N. *et al.* Compromised CDK1 activity sensitizes BRCA-proficient cancers to PARP inhibition. *Nature Med.* **17,** 875–882 (2011).
47. Akamatsu, Y. & Jasin, M. Role for the mammalian Swi5-Sfr1 complex in DNA strand break repair through homologous recombination. *PLoS Genet.* **6,** e1001160 (2010).
48. Mendes-Pereira, A. M. *et al.* Synthetic lethal targeting of *PTEN* mutant cells with PARP inhibitors. *EMBO Mol. Med.* **1,** 315–322 (2009).
49. Brenner, J. C. *et al.* Mechanistic rationale for inhibition of poly(ADP-ribose) polymerase in *ETS* gene fusion-positive prostate cancer. *Cancer Cell* **19,** 664–678 (2011).

**An important paper suggesting that *ETS* gene fusions may confer PARP-inhibitor sensitivity in prostate cancer.**

50. Edwards, S. L. *et al.* Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* **451,** 1111–1115 (2008).
51. Sakai, W. *et al.* Secondary mutations as a mechanism of cisplatin resistance in *BRCA2*-mutated cancers. *Nature* **451,** 1116–1120 (2008).
**References 50 and 51 are the first demonstrations of synthetic lethal resistance to PARP inhibitors and platinum drugs in *BRCA*-mutant cells.**
52. Sakai, W. *et al.* Functional restoration of BRCA2 protein by secondary *BRCA2* mutations in *BRCA2*-mutated ovarian carcinoma. *Cancer Res.* **69,** 6381–6386 (2009).
53. Swisher, E. M. *et al.* Secondary *BRCA1* mutations in *BRCA1*-mutated ovarian carcinomas with platinum resistance. *Cancer Res.* **68,** 2581–2586 (2008).
54. Norquist, B. *et al.* Secondary somatic mutations restoring *BRCA1/2* predict chemotherapy resistance in hereditary ovarian carcinomas. *J. Clin. Oncol.* **29,** 3008–3015 (2011).
55. Rottenberg, S. *et al.* High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. *Proc. Natl Acad. Sci. USA* **105,** 17079–17084 (2008).
56. Bouwman, P. *et al.* 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and *BRCA*-mutated breast cancers. *Nature Struct. Mol. Biol.* **17,** 688–695 (2010).
57. Bunting, S. F. *et al.* 53BP1 inhibits homologous recombination in *Brca1*-deficient cells by blocking resection of DNA breaks. *Cell* **141,** 243–254 (2010).
58. Jiang, H. *et al.* The combined status of ATM and *p53* link tumor development with therapeutic response. *Genes Dev.* **23,** 1895–1909 (2009).
59. Galanty, Y. *et al.* Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* **462,** 935–939 (2009).
60. Soucy, T. A. *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458,** 732–736 (2009).
**This paper provides proof of the concept that targeting additional elements of the DDR and replication machinery, such as neddylation complexes, could provide an alternative approach to cancer therapy.**
61. Lansdorp, P. M. Immortal strands? Give me a break. *Cell* **129,** 1244–1247 (2007).
62. Diehn, M. *et al.* Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* **458,** 780–783 (2009).
63. Zhang, M., Atkinson, R. L. & Rosen, J. M. Selective targeting of radiation-resistant tumor-initiating cells. *Proc. Natl Acad. Sci. USA* **107,** 3522–3527 (2010).
64. Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432,** 396–401 (2004).
65. Bao, S. *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444,** 756–760 (2006).
66. Mantoni, T. S., Lunardi, S., Al-Assar, O., Masamune, A. & Brunner, T. B. Pancreatic stellate cells radioprotect pancreatic cancer cells through beta 1-integrin signaling. *Cancer Res.* **71,** 3453–3458 (2011).
67. Stratton, M. R. Exploring the genomes of cancer cells: progress and promise. *Science* **331,** 1553–1558 (2011).
68. Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462,** 1005–1010 (2009).
69. Barber, L. J. *et al.* Comprehensive genomic analysis of a *BRCA2* deficient human pancreatic cancer. *PLoS ONE* **6,** e21639 (2011).
70. Graeser, M. *et al.* A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. *Clin. Cancer Res.* **16,** 6159–6168 (2010).
71. Polo, S. E. & Jackson, S. P. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev.* **25,** 409–433 (2011).
72. Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411,** 366–374 (2001).
73. Liu, P. *et al.* Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* **146,** 889–903 (2011).

**Acknowledgements** We thank Cancer Research UK, The Wellcome Trust, Breakthrough Breast Cancer, AACR, The Komen Foundation, The Breast Cancer Research Foundation, the European Union and The Breast Cancer Campaign for funding the work in our laboratory. We also acknowledge NHS funding to the Royal Marsden Hospital NIHR Biomedical Research Centre.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to C.J.L. (chris.lord@icr.ac.uk) or A.A. (alan.ashworth@icr.ac.uk)

doi:10.1038/nature10761

# The promise of induced pluripotent stem cells in research and therapy

Daisy A. Robinton[1-5] & George Q. Daley[1-5]

**The field of stem-cell biology has been catapulted forward by the startling development of reprogramming technology. The ability to restore pluripotency to somatic cells through the ectopic co-expression of reprogramming factors has created powerful new opportunities for modelling human diseases and offers hope for personalized regenerative cell therapies. While the field is racing ahead, some researchers are pausing to evaluate whether induced pluripotent stem cells are indeed the true equivalents of embryonic stem cells and whether subtle differences between these types of cell might affect their research applications and therapeutic potential.**

After a decade of constraints, pluripotent stem-cell biology is now a flourishing research area, following the achievement of a long-standing ambition — the successful derivation of pluripotent stem cells from a patient's cells. In a momentous contribution, in 2006 Takahashi and Yamanaka illustrated how cell fates can be altered by the ectopic co-expression of transcription factors[1]. The manipulation of cell fates through reprogramming has altered fundamental ideas about the stability of cellular identity, stimulating major new directions in research into human disease modelling, tissue differentiation *in vitro* and cellular transdifferentiation. Despite heady progress, a major question remains: are the new induced pluripotent stem (iPS) cells equivalent to the classic embryonic stem (ES) cells and thus a suitable alternative for research and therapy? Whereas the initial wave of papers argued convincingly that the two cell types were functionally equivalent, a more refined analysis of how iPS cells behave *in vitro*, coupled with genome-wide genetic and epigenetic analysis, has revealed numerous subtle but substantial molecular differences, probably owing to technical limitations inherent in reprogramming. In this Review, we describe the derivation of iPS cells, outline the functional assessments of pluripotency, and then recount how global assessments of gene expression, gene copy number variation, DNA methylation and chromatin modification provide a more nuanced comparison of iPS cells and ES cells. We detail how these features influence the utility of each of these cell types for disease modelling and therapeutics, and offer predictions for the evolution of the art of reprogramming somatic cells.

## Pluripotent stem cells

The years since Takahashi and Yamanaka's breakthrough have seen a flood of papers touting advances in reprogramming technology, including alternative methods for reprogramming and the successful derivation of iPS cells from various cell types. Although the field has advanced at a breathtaking pace, investigators have recently taken a step back to more critically evaluate iPS cells relative to ES cells and have endeavoured to fully understand how these cell populations differ from one another in the hope of closing the gap between the two populations. Taking clues from the data, it seems that researchers should attempt to define each cell type more accurately and to understand its inherent properties rather than ask whether these two classes of pluripotent cell are identical. Although ES cells and iPS cells are arguably equivalent in all their functions, these cells are bound to harbour subtle differences and to have distinct but complementary roles in research because of their distinct origins and modes of derivation. To appreciate the differences between ES cells and iPS cells, we must first define what it means to be pluripotent.

The term pluripotency has been assigned to a variety of cell types with a wide range of functional capacities. In its loosest sense, pluripotent describes a cell that can generate cell types from each of the three embryonic germ layers: the endoderm, mesoderm and ectoderm. At the strict end of the range of definitions, however, pluripotent describes a cell that can give rise to an entire organism, generating every cell type within that organism. The property of cell pluripotency was first exposed by Driesch in 1891, when he separated the two cells of an early sea urchin blastocyst and observed the development of two complete sea urchins[2]. Many decades later, studies of embryo aggregation and blastocyst chimaerism by Mintz and colleagues[3], Gardner[4] and Brinster[5], in the 1960s and 1970s, solidified the idea that the cells of the inner cell mass of the mouse blastocyst were pluripotent, and the isolation of mouse teratocarcinoma stem cells and native ES cells by Evans and Kaufman[6] and Martin[7], in 1981, ushered in the era of culturing pluripotent stem cells in a dish. The first successful isolation of human ES cells, by Thomson and colleagues in 1998, brought forth a surge of excitement in the scientific community and beyond[8]. The potential to understand early human development, tissue formation and differentiation *in vitro* through studying ES cells seemed to offer limitless possibilities. The opportunity to model diseases, discover disease mechanisms and, ultimately, use cell therapy for previously untreatable conditions was particularly alluring.

The derivation of ES cells from the human embryo, however, sparked controversy in the United States and led to a presidential executive order that restricted government funding[9]. The limited numbers of stem cell lines that were approved for research lacked the diversity necessary to address some of the most compelling questions, particularly those related to modelling and treating disease[10]. Most ES cells represented generic cells isolated from presumably normal embryos — except for those from embryos that had been tested by pre-implantation diagnostics and found to carry genetic

---

[1]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Children's Hospital Boston and Dana Farber Cancer Institute, 300 Longwood Avenue, Boston, Massachusetts 02115, USA. [2]Division of Hematology, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115, USA. [3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. [4]Broad Institute, Cambridge, Massachusetts 02142, USA. [5]Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA.

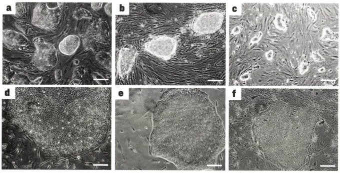


**Figure 1 | Morphology of pluripotent stem cell types.** Mouse ES (**a**) and iPS (**b**) cells form dome-shaped, refractile colonies. These colonies are in contrast to the flat morphology of mouse epiblast-derived stem cells (**f**), which resemble human ES (**d**) and iPS (**e**) cells. Human iPS cells induced into a naive pluripotent state by treatment with chemical inhibitors[97–100] (**c**) parallel the morphology of mouse ES and iPS cells. Scale bars, 50 μm.

diseases. The generic lines were not matched to a particular patient, so products derived from them for transplantation purposes would face rejection by the transplant recipient's immune system or necessitate that the recipient receive lifelong therapy with toxic immunosuppressive medication. To compound these limitations, when human ES cells are cultivated on mouse feeder cells, the human cells can incorporate mouse components that render the ES cells subject to immune rejection.

To realize the full potential of ES cells, researchers foresaw that customized, personalized pluripotent stem cells specific to each patient would be generated by using somatic-cell nuclear transfer (SCNT) — the procedure that had been used successfully to clone Dolly the sheep from adult mammary cells. Nuclear-transfer-generated ES (ntES) cell lines would capture a patient's complete genome in a cell that could be induced to become any tissue, thus allowing differentiation into disease-relevant cells for analysis or cell-replacement therapy. Despite successful proof of principle in mouse studies[11], and the clear distinctions between generating ntES cells for medical research and creating cloned blastocysts for reproduction, the ethical controversy driven by widespread opposition to human cloning has severely curtailed research into human SCNT. Only this year, when investigators gained access to a large number of human oocytes, was the derivation of pluripotent stem cells through human SCNT accomplished[12]. However, the investigators in this study needed to leave the oocyte nucleus intact to derive pluripotent stem cells, so the resultant cells were triploid, thus affording research applications for these cells but limiting their suitability for therapeutic use[12].

Despite the many hindrances to the study and derivation of human ES and ntES cells over the past decade, great strides were being made in understanding the pathways that regulate the maintenance and pluripotency of ES cells. This progress was not lost on those seeking an alternative source of personalized patient-specific stem cells, and in 2006 Takahashi and Yamanaka announced the successful derivation of iPS cells from adult mouse fibroblasts through the ectopic co-expression of only four genes[1]. In an elegant screen of 24 gene candidates selected for their links to ES-cell pluripotency, these researchers found four factors that were sufficient to reprogram adult fibroblasts into iPS cells: OCT4 (also known as POU5F1), SOX2, Krüppel-like factor 4 (KLF4) and c-MYC. This historic contribution inspired an astonishing flurry of follow-up studies, with successful reprogramming quickly translated to human fibroblasts[13–15] and then to a wide variety of other cell types, including pancreatic β cells[16], neural stem cells[17,18], mature B cells[19], stomach and liver cells[20], melanocytes[21], adipose stem cells[22] and keratinocytes[23], demonstrating the seemingly universal capacity to alter cellular identity.

Mouse and human iPS cells differ in appearance. Mouse iPS-cell colonies appear more dome-like and refractile than human iPS-cell colonies. Human iPS-cell colonies are flatter than those of mice and are akin to a distinct type of pluripotent stem cell that is derived from the epiblast of the early mouse embryo[24], a feature that indicates that mouse and human iPS cells, like mouse and human ES cells, probably reflect distinct developmental states (Fig. 1). The pluripotent state of mouse stem cells is called a 'naive' state because it closely resembles the most primitive state, or ground state, of the mouse inner cell mass; this is different from the more 'primed' state of human stem cells, which proliferate in response to different cytokines, reflecting the distinct developmental states of these populations[25]. Regardless of the method of derivation, iPS cells maintain the key features of ES cells, including the ability to propagate in culture indefinitely and the capacity to generate cells from each of the three embryonic germ layers (see ref. 26 for a review). Such broad similarities are not proof that iPS cells are molecularly or functionally equivalent to ES cells. Yamanaka's intention was to derive an alternative source of pluripotent stem cells with the same range of functions as ES cells but offering even greater potential for clinical use. To determine the degree of success garnered by reprogramming, we must explore the set of assays that were developed to assess the key characteristic of ES cells: pluripotency.

## Assessment of pluripotency

In the past few years, consistent standards for the identification and evaluation of iPS cells and for the assessment of their functional equivalence to ES cells have become widely accepted[27]. A variety of reprogramming methods have been developed to derive iPS cells, and each has advantages and disadvantages (Table 1). Assessing reprogramming begins with identifying compact colonies that have distinct borders and well-defined edges, and are comprised of cells with a large nucleus, large nucleoli and scant cytoplasm. A wide range of colony morphologies result from reprogramming, and although many colonies appear morphologically similar to ES cells, only a subset of these have comparable molecular and functional features. To accurately distinguish reprogrammed, bona fide iPS cells from those that are only partially reprogrammed, investigators look for a series of molecular hallmarks.

## Markers of pluripotency

Fully reprogrammed cells express a network of pluripotency genes, including *OCT4*, *SOX2* and *NANOG*, in levels comparable to ES cells, and they reactivate telomerase gene expression, downregulate *THY1* and upregulate *SSEA1* (ref. 28). Positive staining for alkaline phosphatase activity has been widely used as a marker of pluripotency; however, recently published data have shown this to be insufficient as a test for true iPS cells, because intermediately reprogrammed cells also stain positively[29]. The same report shows that iPS cells that are generated by virus-mediated reprogramming silence proviral genes when the endogenous pluripotency genes are activated, and that this event is paired with the expression of the embryonic antigens SSEA3, TRA-1-60, TRA-1-81, DNA methyltransferase 3β (DNMT3β) and REX1 (ref. 29). Genome-wide epigenetic reprogramming is crucial for deriving fully reprogrammed cells, and the degree of success is measured, in part, by evaluating the methylation status at the promoters of the genes responsible for maintaining pluripotency, as well as at the genes important for driving differentiation[30]. A crucial event during epigenetic reprogramming is the reactivation of the silent X chromosome, which occurs late in reprogramming and represents a hallmark of ground-state pluripotency[28,30,31]. If iPS cells acquire all of these molecular features, they are expected to behave like ES cells and to demonstrate reprogramming-factor independence, which is marked by silencing of the proviral transgenes. Variations in epigenetic reprogramming, the extent of methylation, the persistence of expression of integrated proviruses and other known and unknown factors can alter the differentiation potential of iPS cells. Because of the potential for heterogeneity, it is essential to know as much as possible about the nature of a cell line before labelling it pluripotent.

**Table 1 | Methods for reprogramming somatic cells to iPS cells**

| Vector type | | Cell types | Factors* | Efficiency (%) | Advantages | Disadvantages |
|---|---|---|---|---|---|---|
| Integrating | Retroviral[1,14,82,83] | Fibroblasts, neural stem cells, stomach cells, liver cells, keratinocytes, amniotic cells, blood cells and adipose cells | OSKM, OSK, OSK + VPA, or OS + VPA | ~0.001–1 | Reasonably efficient | Genomic integration, incomplete proviral silencing and slow kinetics |
| | Lentiviral[15,16,84,85] | Fibroblasts and keratinocytes | OSKM or *miR302/367* cluster + VPA | ~0.1–1.1 | Reasonably efficient and transduces dividing and non-dividing cells | Genomic integration and incomplete proviral silencing |
| | Inducible lentiviral[23,28] | Fibroblasts, $\beta$ cells, keratinocytes, blood cells and melanocytes | OSKM or OSKMN | ~0.1–2 | Reasonably efficient and allows controlled expression of factors | Genomic integration and requirement for transactivator expression |
| Excisable | Transposon[86] | Fibroblasts | OSKM | ~0.1 | Reasonably efficient and no genomic integration | Labour-intensive screening of excised lines |
| | *loxP*-flanked lentiviral[87] | Fibroblasts | OSK | ~0.1–1 | Reasonably efficient and no genomic integration | Labour-intensive screening of excised lines, and *loxP* sites retained in the genome |
| Non-integrating | Adenoviral[88,89] | Fibroblasts and liver cells | OSKM | ~0.001 | No genomic integration | Low efficiency |
| | Plasmid[90,91] | Fibroblasts | OSNL | ~0.001 | Only occasional genomic integration | Low efficiency and occasional vector genomic integration |
| DNA free | Sendai virus[92] | Fibroblasts | OSKM | ~1 | No genomic integration | Sequence-sensitive RNA replicase, and difficulty in purging cells of replicating virus |
| | Protein[93,94] | Fibroblasts | OS | ~0.001 | No genomic integration, direct delivery of transcription factors and no DNA-related complications | Low efficiency, short half-life, and requirement for large quantities of pure proteins and multiple applications of protein |
| | Modified mRNA[95] | Fibroblasts | OSKM or OSKML + VPA | ~1–4.4 | No genomic integration, bypasses innate antiviral response, faster reprogramming kinetics, controllable and high efficiency | Requirement for multiple rounds of transfection |
| | MicroRNA[96] | Adipose stromal cells and dermal fibroblasts | miR-200c, miR-302s or miR-369s | ~0.1 | Efficient, faster reprogramming kinetics than commonly used lentiviral or retroviral vectors, no exogenous transcription factors and no risk of integration | Lower efficiency than other commonly used methods |

*OSKM and similar factor names represent combinations of reprogramming factors: K, KLF4; L, LIN28; M, c-MYC; N, NANOG; O, OCT4; S, SOX2; and VPA, valproic acid.

## Functional assays of pluripotency

When iPS cell lines are isolated and documented to carry the molecular features of fully reprogrammed cells, they are typically also assessed in functional assays. Characterization of the functional abilities of iPS cells begins with *in vitro* differentiation. The cells can be differentiated as embryoid bodies — compact balls of loosely organized tissues that resemble the gastrulating embryo — or through two-dimensional directed differentiation in a culture dish. Such cultures can then be assessed for markers of each of the three germ layers. Analysis of the pluripotency of mouse cells typically entails the development of a chimaera, which evaluates the potential of iPS cells to contribute to the normal development of adult tissues after injection into the blastocyst. Whether germline transmission occurs after blastocyst chimaerism is measured by the ability of chimaeras to produce all-iPS-cell mice in their offspring. These offspring have the genomic integrity of the injected iPS cell line, as well as the ability to form functional germ cells. The highest stringency test for mouse iPS cells — tetraploid complementation — entails the injection of iPS cells into tetraploid blastocysts to measure the ability of the iPS cells to direct the normal development of an entire organism. This test has been accomplished for only a limited subset of iPS cells[32–34], although with an efficiency that parallels tetraploid complementation carried out with ES or ntES cells[33,35,36].

The current functional gold standard for human iPS cells involves the evaluation of teratoma formation. In this assay, the *in vivo* differentiation potential of human iPS cells is measured after their injection subcutaneously or intramuscularly into immunodeficient mice[37,38]. If the cells are truly pluripotent, they will form well-differentiated tumours comprised of elements from each of the three germ layers. This assay provides information about the spontaneous differentiation potential of the injected iPS cells. Although it is the most stringent assay available for human iPS cells, it is not powerful enough to assess whether iPS cells can produce all the cell types of the human body, and it cannot assess the contribution of iPS cells to the germ line. The caveat to all these functional assays lies in the fact that the standards for iPS cells are still hotly debated, especially when anticipating the use of iPS cells for therapy[39]. Adopting a consistent set of standards that can be applied uniformly worldwide is essential as stem-cell research and applications move forward.

## Functional differences between iPS cells and ES cells

Despite the multitude of assays used to evaluate pluripotency, and although many parallels have been found between iPS cells and ES cells, there is a wide range of evidence showing that there are subtle yet substantial differences between these cell types. Disparities were first observed in the differentiation abilities of iPS cells and ES cells in both teratoma-forming and *in vitro* differentiation assays. Some mouse iPS cells showed lower efficiencies of teratoma formation than mouse ES cells, whereas some human iPS cells showed less propensity to differentiate along haematopoietic, neuroepithelial and neuronal lineages than human ES cells[40–42]. Some researchers interpreted these findings to mean that iPS cells have an intrinsically lower differentiation capacity than ES cells[41], whereas other research groups have offered different explanations, including that the cell of origin might have a specific effect on the differentiation capacities of the derived iPS cells.

The results from cell-of-origin studies indicate that the parental cell can influence the differentiation capacity of the resultant iPS cells. In one study, mouse bone-marrow-derived and B-cell-derived iPS cells showed more efficient differentiation along haematopoietic lineages than did fibroblast-derived iPS cells or neural-progenitor-derived iPS cell lines. Interestingly, treatment of the neural-progenitor-derived iPS cells with trichostatin A, a potent histone-deacetylase inhibitor, plus 5-azacytidine, a methylation-resistant cytosine analogue, increased the blood-forming capacity of these cells, suggesting that their limitations were due to epigenetic modifications. Whereas the bone-marrow-derived and neural-progenitor-derived iPS cells contributed well to all tissues in the chimaera assay, including to the germ line, the fibroblast-derived iPS cells contributed only poorly[43]. This study laid some of the early groundwork for later lines of investigation that probed the molecular differences between iPS cells and ES cells, and provided an explanation for the functional differences between these cells.

One investigation found that some iPS cells derived from human retinal-pigment epithelial cells show an increased propensity to differentiate back into this cell type than do ES cells or iPS cells derived from other tissues[44]. More recently, Bar-Nur and colleagues showed that iPS cells generated from human pancreatic islet β cells retain open chromatin at the loci of key β-cell genes and that this correlates with a greater capacity to differentiate into insulin-producing cells both *in vivo* and *in vitro* than that of ES cells or isogenic non-β-cell-derived iPS cells[45]. These functional differences extend beyond differentiation and potency to disease modelling. For example, fragile X syndrome is caused by aberrant silencing of the *FMR1* gene during human development; iPS cells that were reprogrammed from adult skin fibroblasts from an individual with fragile X syndrome failed to reactivate the *FMR1* gene, whereas ES cells derived from embryos with this syndrome, as diagnosed by pre-implantation testing, expressed *FMR1* (ref. 46). Consequently, the potential for epigenetic memory in the fragile-X-syndrome-derived iPS cells, and substantial differences between fragile-X-syndrome-derived iPS and ES cells, must be considered when studying this condition and potentially many other conditions. To determine whether the pluripotent cells being used are appropriate to address a particular question or to use in a given application, it is crucial to compare not only the *in vivo* and *in vitro* differentiation potentials but also the genetic and epigenetic disparities that underpin these functional differences.

## iPS cells versus ES cells

Refined analyses, described in this section, have addressed whether iPS cells are suitable alternatives to classic ES cells for use in research and therapy.

### Genetics and epigenetics

Global gene-expression analysis and bisulphite genomic sequencing accompanying early derivations of iPS cells provided the initial evidence for differences between iPS cells and ES cells at the epigenetic level[14]. Further exploration of these differences led to the identification of only a few, seemingly consistent, differences in global gene expression that were more pronounced in earlier passages of iPS cells[47]. Many of the differentially expressed genes were imprinted in ES cells[48].

Looking beyond the expression patterns to the DNA sequence itself has revealed genetic variation between iPS cells and ES cells. A recent publication suggested that chromosomal aberrations are a common feature of stem-cell populations that are propagated *in vitro*, with each type of stem cell — whether ES cell, iPS cell or multipotent stem cell — being prone to distinct abnormalities[49]. Both human iPS and ES cells showed a tendency for gains at chromosomes 12 and 17. Whereas iPS cells had additional gains at chromosomes 1 and 9, ES cells had additional gains at chromosomes 3 and 20. Other work identified an accumulation of point mutations in reprogrammed cells, particularly occurring in oncogenic pathways[45], whereas another study noted an increase in copy number variants (CNVs) in early-passage human iPS cells relative to intermediate-passage iPS cells or ES cells[50]. The number and size of these CNVs were negatively correlated with the passage number in iPS cells, suggesting that a selective disadvantage is conferred by these aberrations. A comparison of iPS cells and their parental cell of origin showed that the majority of CNVs were created *de novo* in fragile regions of the genome[50]. A comprehensive study by Laurent and colleagues found a higher frequency of subchromosomal CNVs in pluripotent cell samples than in non-pluripotent cell samples[51]. This work uncovered variation between genomic regions enriched for CNVs in human ES cells and iPS cells. A small subset of samples of ES cells harboured a large number of duplications, whereas several iPS-cell samples contained moderate numbers of deletions. Reprogramming was associated with deletions in tumour-suppressor genes, whereas extended time in culture led to duplications of oncogenes in human iPS cells.

The finding that human iPS cells derived from a variety of tissues have residual, persistent donor-cell-specific gene-expression patterns sparks the question of whether the current measure of a fully reprogrammed cell is sufficient[52] or whether iPS cells retain some type of 'somatic memory' from their past identity. To understand this observation better, a more detailed analysis at the epigenetic level is required.

Reprogramming cells to a pluripotent state entails global epigenetic remodelling and introduces epigenetic changes, some of which are necessary for reprogramming to occur and others of which are inadvertently introduced during the process. A failure to demethylate pluripotency genes is associated with partial reprogramming in iPS cells. Whole-genome profiling of the DNA methylomes of multiple human iPS and ES cell lines, as well as somatic and progenitor cell lines, from different laboratories using different reprogramming techniques and with a variety of cells derived from distinct germ layers has shown that although overall iPS-cell DNA methylomes closely resemble human ES-cell DNA methylomes, iPS cells have significant variability in their somatic memory, as well as aberrant iPS-cell-specific differential methylation. Some studies have suggested that this occurs in a passage-dependent manner, but others have shown that differentially methylated regions (DMRs) in iPS cells are transmitted to differentiated progeny at a high frequency and cannot be erased through passaging[29,53–55]. Overall, there are remarkable global similarities between the DNA methylomes of generic iPS and ES cells; however, a core set of DMRs that seems to represent hot spots of failed epigenetic reprogramming has been identified[55]. These DMRs are enriched for genes that are important for developmental processes[51,55]. The high incidence of unique DMRs in iPS cells compared with progenitor somatic cells or ES cells suggests that these patterns are stochastic and arise during reprogramming. In the most exhaustive comparison so far, Kim and colleagues reported that more DMRs were present in mouse iPS cells than in ntES or embryo-derived ES cells[43]. However, these DMRs did not pertain to specific loci and thus do not represent consistent differences between iPS cells and ES cells. This lack of consistency suggests that aberrant DMRs in mouse iPS cells reflect the technical limitations inherent in reprogramming, rather than indicating loci that can reliably distinguish ES cells from iPS cells[43].

In addition, the residual iPS-cell-specific methylation in many iPS-cell isolates links these cells to their tissue of origin and, ultimately, affects their differentiation propensity[43,55]. Residual signatures can be distinct enough to enable the myeloid and lymphoid origins of blood-derived iPS cells to be discerned[43]. In iPS cells derived from non-haematopoietic cells, such as fibroblasts and neural progenitors, there can be residual repressive methylation at loci that are required for haematopoietic fates, reducing the blood-forming potential *in vitro*[43]. Exogenous supplementation of neural-progenitor-derived iPS cells with the cytokine WNT3A can increase the blood-forming potential of these cells, supporting the idea that incomplete reprogramming owing to epigenetic marks can be overcome by manipulating the culture conditions. Treatment of cultures with demethylating agents or knockdown of *DNMT1* expression has been shown to convert intermediately reprogrammed cells into fully pluripotent cells, further supporting this idea[48]. When iPS cells are forced to differentiate along a particular lineage, they become more amenable to generating cells of that lineage after another round of reprogramming[43]. This finding shows that the differentiation propensity and DNA methylation profile can be reset, and it suggests that the 'epigenetic memory' of the donor cell can be exploited, especially in cases in which directed differentiation is particularly challenging[43].

Another important feature of epigenetic reprogramming is the reactivation of the inactivated X chromosome. During normal development in eutherian mammals (those with a placenta), one X chromosome is randomly inactivated in each cell in females. Whether this epigenetic silencing event is reset in iPS cells remains an area of controversy, in part because of the poor fidelity of X-inactivation markers in pluripotent cells[56]. Some studies have shown that the majority of female human iPS clones retain an inactivated X chromosome (which is transcriptionally silent)[57], whereas others have indicated that some human iPS clones lose immunostaining for trimethylated H3K27 on the X chromosome (a marker of epigenetic silencing), indicating X reactivation[58]. In addition, some of the earliest studies of iPS cells showed X reactivation in reprogrammed female mouse fibroblasts[31,58]. However, recently published data support the finding that X reactivation does not occur in human iPS cells and, interestingly, reprogramming was found to favour expression of a particular X chromosome when induced from a mixed X-inactivated population of fibroblasts[59]. Finally, epigenetic reprogramming sometimes fails to properly restore bivalent domains, which mark developmental loci with active and repressive histone modifications[60].

Although many of the studies cited here have generated data suggesting that there are epigenetic differences between iPS cells and ES cells, there are several limitations on extending these data to all iPS cells in a more general (and more useful) sense. The published comparisons were often made using iPS cells derived from a multitude of laboratories by a variety of methodologies, and reanalysis of the gene-expression microarray data using an unsupervised clustering algorithm shows a strong correlation between transcriptional signatures and specific laboratories for both iPS cells and ES cells. This finding indicates that specific culture protocols and laboratory environments can affect the transcriptional profile of iPS and ES cells. Therefore, the data produced in a particular laboratory might be specific to the cells derived there[61].

In addition, most iPS colonies are clones derived from a single reprogrammed cell, whereas ES cells used for analysis are typically nonclonal. The subcloning of ES cells has revealed genetic and epigenetic anomalies that would probably have otherwise gone undetected in the heterogeneous ES-cell population[62]. With regard to somatic memory, there is poor overlap between the gene sets that have been reported to be characteristic of a particular cell type of origin, suggesting that the retention of somatic memory is stochastic and is a reflection of the technical failure of reprogramming to fully erase the somatic epigenome. To exacerbate the issue, the iPS cells used for comparison often have different genetic backgrounds and have frequently been derived from fibroblasts that were already heterogeneous in their make-up, affecting both the gene-expression patterns and the functionality of the iPS cells.

Throughout the literature, many publications lack correlation between the gene-expression patterns and the epigenetic patterns observed. An additional consideration is the presence of different viral insertions in individual iPS cell lines, which can also affect the functionality of the derived cells[1]. Evidence to support this idea is provided by the reduced number of differences observed among iPS cells and between iPS and ES cells when transgenes are removed[63]. Many of the aforementioned studies have focused on differences in either transcriptional profiles or changes in epigenetic marks; however, the most recent studies have evaluated iPS cells and ES cells from both of these angles in parallel, together with their *in vitro* differentiation potential, generating the most comprehensive and compelling data that have been published so far.

## Holistic analysis

Stadtfeld and colleagues explored the epigenetic and functional discrepancies between iPS cells and ES cells using a new reprogramming strategy that allowed direct comparison of genetically matched cells derived from the same source[32]. These authors derived iPS and ES cells from mice carrying an integrated doxycycline-inducible reprogramming cassette in every cell, a strategy that sidesteps the confounding effects of variable genetic backgrounds and viral integration that have been observed in other studies. The overall messenger RNA and microRNA expression patterns of iPS cells and ES cells were indistinguishable except for the aberrant silencing of a few transcripts localized to the imprinted *Dlk1–Dio3* gene cluster on chromosome 12qF1, a region that is important for development. A failure to reactivate this locus meant that iPS cells contributed poorly to chimaeras and were unable to generate all-iPS-cell mice. By contrast, iPS cells with normal *Dlk1–Dio3* expression contributed to high-grade chimaeras and supported the development of viable all-iPS-cell mice. The treatment of iPS cells that failed to reactivate *Dlk1–Dio3* with a histone-deacetylase inhibitor rescued the ability of these clones to support the development of all-iPS-cell mice by relieving this region of aberrant hypermethylation. However, recent data from iPS and ES cells derived from a mouse strain carrying a distinct reprogramming cassette suggest that different levels of expression of reprogramming factors, rather than aberrant silencing of *Dlk1–Dio3*, account for the different behaviour of the cell lines[64]. The disparate results from these studies highlight that iPS cells can behave differently based on subtle variations in the expression of only a few loci.

To systematically compare human iPS cells derived from different somatic cell types and ES cells, Ohi and colleagues compared ES cells with iPS cells reprogrammed from somatic cells representative of the three embryonic germ layers[65]. Transcriptional and epigenetic profiling of these cells showed transcriptional differences, owing, in part, to incomplete promoter methylation, which enabled iPS cells to be discerned on the basis of their cell of origin. The differential methylation between iPS cells and ES cells did not correlate with varying levels in DNA methyltransferases; however, the authors found a nonrandom pattern of incompletely silenced genes in genetic regions that are isolated from other genes that undergo silencing during reprogramming. This finding could be explained by inefficient or delayed recruitment of the silencing machinery and DNA methyltransferases to particular somatic genes because of the isolated location of these genes[65].

In another comprehensive study, Polo and colleagues evaluated the effect of cellular origin on the gene-expression pattern, epigenetic properties and functional abilities of genetically matched mouse iPS cells[66]. Using the same 'secondary' reprogramming strategy used by Wernig and colleagues[53], whereby reprogramming is assessed using tissues from a mouse generated from iPS cells carrying integrated, doxycycline-inducible reprogramming factors, Polo and colleagues generated iPS cells from tail-tip fibroblasts, splenic B cells, bone-marrow-derived granulocytes and skeletal muscle precursors, and showed that each iPS cell line retained a transcriptional memory of its cell of origin. This memory was evident in that markers for each respective cell of origin remained actively expressed, and was supported by the finding that iPS cell lines derived from the same cell of origin clustered together on the basis of global transcriptional data. A similar correlation was found on evaluation of methylation patterns, which showed subtle but substantial differences, reflecting the consequences of different histone marks. The effects of somatic memory extended beyond the genetic and epigenetic levels to functional significance, affecting the autonomous differentiation potential of the different iPS cell lines after embryoid-body formation. A clear bias that reflected the cells of origin was observed in the iPS cell lines. Notably, the transcriptional, epigenetic and functional effects evaluated in early-passage iPS cell lines became less significant with continued passaging. This finding indicated that complete reprogramming is a gradual process that extends beyond the time frame necessary to observe the activation of endogenous pluripotency genes, transgene-free growth and differentiation into cell types from each of the three germ layers.

### Exploring equivalence

Having considered data from a multitude of published studies, generated by the painstaking efforts of many research groups, we return to our earlier question: are iPS cells equivalent to ES cells? The answer is not straightforward. Rather, there is an emerging consensus that iPS cells and ES cells are neither identical nor distinct populations. Instead, they are overlapping, with greater variability inherent within each population than between the populations. The heterogeneity and behaviour of each class of cells is more complex than has previously been thought. The two pluripotent stem-cell types are, in theory, functionally equivalent; however, in practice, they harbour genetic and epigenetic differences that reflect their different histories. It remains to be seen whether there are any consistent molecular distinctions between iPS cells and ES cells.

It is also important to consider that, in contrast to long-standing belief, ES cells themselves have considerable epigenetic heterogeneity and have differing propensities for differentiation — much like those found in iPS cells[67,68]. These observations, paired with those discussed in this Review, are a call for researchers to take a step back from the direct comparison of iPS cells and ES cells, and they highlight the need to redefine what it means to belong to either of these cell classes. Some researchers have already taken heed of this message and generated a bioinformatics assay for pluripotency[63], whereas others have produced a 'scorecard' to evaluate the character of both iPS and ES cell lines and predict the quality and utility of any pluripotent cell line in a high-throughput manner[47,69]. Using DNA methylation mapping, gene-expression profiling and a quantitative differentiation assay, Bock and colleagues made a systematic comparison of 20 established ES cell lines and 12 iPS cell lines[47]. They confirmed that, despite overall similarities, transcriptional and epigenetic variation is common between iPS cell lines, between ES cell lines, and between iPS cell and ES cell populations. These data provide a reference for the variation among human pluripotent cell lines, which assists in predicting the functional consequences of these differences. We can conclude from these studies that any given iPS cell line generated by today's technology might not be completely equivalent to the ideal ES cell.

The differences between iPS cells and ES cells, as well as those among iPS cells, clearly affect the utility of these cells in research, disease modelling and therapeutics, providing an impetus for investigators to evaluate their cell populations carefully and precisely. The differences do not diminish the potential of iPS cells, given that iPS cells have considerable advantages over ES cells. Rather than replacing ES cells with iPS cells, it is becoming clear that these two cell types complement one another. Researchers are still in the process of developing the necessary protocols to harness the potential of iPS cells; however, as it becomes clear how to evaluate the genetic, epigenetic and functional status of different iPS cell lines, further applications of these cells will be uncovered, and progress will be made in creating iPS cell lines and designing protocols to accomplish the ambitious goals of the field.

## Medical applications of iPS cells

Generating patient-specific stem cells has been a long-standing goal in the field of regenerative medicine. Despite considerable challenges, generating disease-specific and patient-specific iPS cells through reprogramming has become almost routine. These cells provide a unique platform from which to gain mechanistic insight into a variety of diseases, to carry out *in vitro* drug screening, to evaluate potential therapeutics and to explore gene repair coupled with cell-replacement therapy (Fig. 2). In the past few years, the number of reports on applications of iPS cells has steadily increased, testifying to the broad influence of this breakthrough technology (Table 2). Despite the continued presence of substantial hurdles, the pace of this work is such that no review can capture the current state of the field; thus, we point to a few publications that highlight the promising medical applications of iPS cells but also indicate their key limitations.

In 2009, Lee and colleagues harnessed iPS cells to demonstrate disease modelling and drug screening for familial dysautonomia, a rare genetic disorder of the peripheral nervous system[70]. In almost all cases, familial dysautonomia is caused by a single point mutation in the gene encoding the inhibitor of nuclear factor-κB (IκB)-kinase-complex-associated protein (*IKBKAP*) that manifests as an extensive autonomic nervous system deficit and dysfunction in small-fibre sensory neurons. Although many traditional cell-based models have been used to study the pathogenesis of familial dysautonomia and to screen for candidate drugs, none has used symptom-relevant human cell types. With the successful derivation of iPS cells from patients with familial dysautonomia, investigators produced central and peripheral nervous system precursors and subsequently found three disease-related phenotypes, thus providing validation that disease-relevant cell types could accurately reflect disease pathogenesis *in vitro*[70]. After screening with multiple compounds, the authors showed that the disease phenotype could be partially normalized by kinetin, a plant hormone. This initial report demonstrated how iPS cells can facilitate the discovery of therapeutic compounds and described how these cells provided a platform for modelling different severities of familial dysautonomia and for generating predictive tests to determine differences in the clinical manifestation of the disorder.

Such applications of iPS cells in drug screening and discovery are destined to expand to encompass numerous disease conditions. Several research groups have generated models of long QT syndrome, a congenital disease with 12 types, each of which is associated with abnormal ion-channel function, a prolonged QT interval on an electrocardiogram and a high risk of sudden cardiac death due to ventricular tachyarrhythmia. Much work has been carried out in animal models to probe the underlying mechanisms of this syndrome, but cardiomyocytes have distinct and complex electrophysiological properties that differ between species. In addition, the lack of *in vitro* sources of human cardiomyocytes and the inability to model patient-specific variations of this disease has impeded studies.

In a proof-of-principle study for using iPS cells to capture the physiological mechanisms of genetic variation, Moretti and colleagues differentiated iPS cells from individuals with type 1 long QT syndrome into cardiomyocytes and, as predicted, observed prolonged action potentials in the ventricular and atrial cells[71]. Using this model system, these investigators uncovered a dominant-negative trafficking defect associated with the particular mutation that causes this variant of long QT syndrome. Further investigation of long QT syndrome iPS-cell-derived cardiomyocytes showed that these cells had an increased susceptibility to catecholamine-induced tachyarrhythmia, and compounds that exacerbated the condition (including isoprenaline) were identified[71]. Treatment of these cardiomyocytes with β-adrenergic receptor blockers attenuated the long QT phenotype.

Type 2 long QT syndrome has also been modelled in cardiomyocytes, by Itzhaki and colleagues[72]. The authors derived type 2 long QT syndrome iPS cells to evaluate the potency of existing and new pharmacological agents that might exacerbate or ameliorate the condition. Their studies show that the long QT syndrome phenotype was aggravated by blockers of ERG-type potassium channels, whereas nifedipine, a calcium-channel blocker, and pinacidil, an agonist of ATP-sensitive potassium channels, both ameliorated the long QT syndrome phenotype, as shown by the decreased duration of action potentials in long QT syndrome cardiomyocytes, as well as the elimination of early after-depolarizations and the abolishment of all triggered arrhythmias. A possible limitation of these beneficial drugs is excessive shortening of the action-potential duration, leading to short QT syndrome.

Importantly, these studies established that the iPS-cell model can be used to identify complex cardiotoxic effects of drugs, as well as to define protective pharmacological agents, including optimal drug dosages.

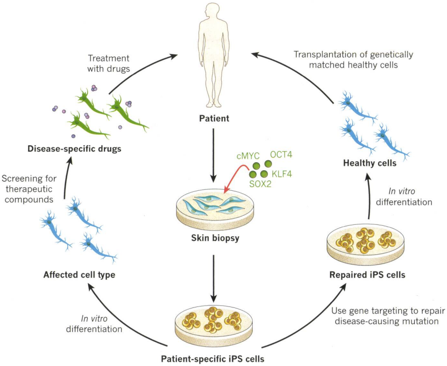


**Figure 2 | Medical applications of iPS cells.** Reprogramming technology and iPS cells have the potential to be used to model and treat human disease. In this example, the patient has a neurodegenerative disorder. Patient-specific iPS cells — in this case derived by ectopic co-expression of transcription factors in cells isolated from a skin biopsy — can be used in one of two pathways. In cases in which the disease-causing mutation is known (for example, familial Parkinson's disease), gene targeting could be used to repair the DNA sequence (right). The gene-corrected patient-specific iPS cells would then undergo directed differentiation into the affected neuronal subtype (for example, midbrain dopaminergic neurons) and be transplanted into the patient's brain (to engraft the nigrostriatal axis). Alternatively, directed differentiation of the patient-specific iPS cells into the affected neuronal subtype (left) will allow the patient's disease to be modelled *in vitro*, and potential drugs can be screened, aiding in the discovery of novel therapeutic compounds.

Given the number of drugs that have notoriously been withdrawn from the market because of their tendency to induce arrhythmias, it is highly likely that the current inadequate approaches for assessing cardiotoxicity will be complemented by iPS-cell-based assessments of drug effects.

A study from our laboratory explored dyskeratosis congenita, a disorder of telomere maintenance, and provided an unanticipated insight into the basic biology of telomerase that has therapeutic implications[73]. In its most severe form, dyskeratosis congenita is caused by a mutation in the dyskerin gene (*DKC1*), which is X linked, leading to shortened telomeres and premature senescence in cells and ultimately manifesting as the degeneration of multiple tissues. Because the reprogramming of cells to an induced pluripotent state is accompanied by the induction of the gene encoding telomerase reverse transcriptase (*TERT*), we investigated whether the telomerase defect would limit the derivation and maintenance of iPS cells from individuals with dyskeratosis congenita. Although the efficiency of iPS-cell derivation was poor, we were able to successfully reprogram patient fibroblasts. Surprisingly, whereas the mean telomere length immediately after reprogramming was shorter than that of the parental fibroblast population, continued passage of some iPS cell lines led to telomere elongation over time. This process was accompanied by upregulation of the expression of *TERC*, which encodes the RNA subunit of telomerase.

Further analysis established that TERT and TERC, as well as DKC1, were expressed at higher levels in dyskeratosis-congenita-derived iPS cells than in the parental fibroblasts[73]. We determined that the genes encoding these components of the telomerase pathway — including a *cis* element in the 3′ region of the *TERC* locus that is essential for a transcriptionally active chromatin structure — were direct binding targets of the pluripotency-associated transcription factors. Further analysis indicated that transcriptional silencing owing to a 3′ deletion in the *TERC* locus leads to the autosomal dominant form of dyskeratosis congenita by diminishing *TERC* transcription. Although telomere length is restored in dyskeratosis-congenita-derived iPS cells, differentiation into somatic cells is accompanied by a return to pathogenesis with low *TERC* expression and a decay in telomere length. This finding showed that TERC RNA levels are dynamically regulated and that the pluripotent state of the cells is reversible, suggesting that drugs that elevate or stabilize *TERC* expression might rescue defective telomerase activity and provide a therapeutic benefit. Although we set out to understand the pathogenesis of dyskeratosis congenita with this study, we showed that a high expression level of multiple telomerase components was characteristic of the pluripotent state more generally, illustrating how iPS cells can reveal fundamental aspects of cell biology.

An independent study of the reprogramming of cells from patients with dyskeratosis congenita confirmed the general transcriptional upregulation of multiple telomerase components and the maintenance of telomere lengths in clones[74]; however, in this study, no clones with elongated telomeres were identified. The different outcomes of these studies highlight the limitations of iPS-cell-based disease models that are imposed by clonal variation as a result of the inherent technical infidelity of reprogramming[75]. This point also introduces an additional important consideration. Before a given iPS-cell disease model can be claimed to be truly representative of the disease, how many patients must be involved, and how many iPS cell lines must be derived from each patient? Although the answers to these questions are unclear, it is crucial to keep these issues in mind when generating disease models and making claims based on these models.

Although iPS cells are an invaluable tool for modelling diseases *in vitro*, the goal of developing patient-specific stem cells has also been motivated by the prospect of generating a ready supply of immune-compatible cells and tissues for autologous transplantation. At present, the clinical translation of iPS-cell-based cell therapies seems more futuristic than the *in vitro* use of iPS cells for research and drug development, but two groundbreaking studies have provided the proof of principle in mouse models that the dream might one day be realized. Hanna, Jaenisch and colleagues used homologous recombination to repair the genetic defect in iPS cells derived from a humanized mouse model of sickle-cell anaemia[76]. Directed differentiation of the repaired iPS cells into haematopoietic progenitors followed by transplantation of these cells into the affected mice led to the rescue of the disease phenotype. The gene-corrected iPS-cell-derived haematopoietic progenitors showed stable engraftment and correction of the disease phenotype.

In another landmark study from Jaenisch's research group, Wernig and colleagues derived dopaminergic neurons from iPS cells that, when implanted into the brain, became functionally integrated and improved the condition of a rat model of Parkinson's disease[77]. The successful implantation and functional recovery in this model is evidence of the therapeutic value of pluripotent stem cells for cell-replacement therapy in the brain — one of the most promising areas for the future of iPS-cell applications.

**Table 2 | Diseases modelled with iPS cells**

| Disease | Molecular defect of donor cell | Cell type differentiated from iPS cells | Disease phenocopied in differentiated cells | Drug or functional tests |
|---|---|---|---|---|
| **Neurological** | | | | |
| Amyotrophic lateral sclerosis (ALS) | Heterozygous Leu144Phe mutation in SOD1 | Motor neurons and glial cells | ND | No |
| Spinal muscular atrophy (SMA) | Mutations in *SMN1* | Neurons and astrocytes, and mature motor neurons | Yes | Yes |
| Parkinson's disease | Multifactorial; mutations in *LRRK2* and/or *SNCA* | Dopaminergic neurons | No | Yes |
| Huntington's disease | 72 CAG repeats in the huntingtin gene | None | NA | No |
| Down's syndrome | Trisomy 21 | Teratoma with tissue from each of the three germ layers | Yes | No |
| Fragile X syndrome | CGG triplet repeat expansion resulting in the silencing of *FMR1* | None | NA | No |
| Familial dysautonomia | Mutation in *IKBKAP* | Central nervous-system lineage, peripheral neurons, haematopoietic cells, endothelial cells and endodermal cells | Yes | Yes |
| Rett's syndrome | Heterozygous mutation in *MECP2* | Neural progenitor cells | Yes | Yes |
| Mucopolysaccharidosis type IIIB (MPS IIIB) | Homozygous mutation in *NAGLU* | Neural stem cells and differentiated neurons | Partially | Yes |
| Schizophrenia | Complex trait | Neurons | Yes | Yes |
| X-linked adrenoleukodystrophy (X-ALD), childhood cerebral ALD (CCALD) and adrenomyeloneuropathy (AMN) | Mutation in *ABCD1* | Oligodendrocytes and neurons | Partially | Yes |
| **Haematological** | | | | |
| *ADA* SCID | Mutation or deletion in *ADA* | None | ND | No |
| Fanconi's anaemia | *FAA* and *FAD2* corrected | Haematopoietic cells | No (corrected) | No |
| Schwachman–Bodian–Diamond syndrome | Multifactorial | None | NA | No |
| Sickle-cell anaemia | Homozygous HbS mutation | None | NA | No |
| β-Thalassaemia | Homozygous deletion in the β-globin gene | Haematopoietic cells | ND | No |
| Polycythaemia vera | Heterozygous Val617Phe mutation in *JAK2* | Haematopoietic progenitors (CD34$^+$CD35$^+$) | Partially | No |
| Primary myelofibrosis | Heterozygous mutation in *JAK2* | None | NA | No |
| **Metabolic** | | | | |
| Lesch–Nyhan syndrome (carrier) | Heterozygous mutation in *HPRT1* | None | NA | No |
| Type 1 diabetes | Multifactorial; unknown | β-Cell-like cells (express somatostatin, glucagon and insulin; glucose-responsive) | ND | No |
| Gaucher's disease, type III | Mutation in *GBA* | None | NA | No |
| α1-Antitrypsin deficiency (A1ATD) | Homozygous mutation in the α1-antitrypsin gene | Hepatocyte-like cells (fetal) | Yes | No |

Together, these findings provide proof of principle for using reprogramming with gene repair and cell-replacement therapy for treating diseases. Using iPS cells in cell-replacement therapy offers the promise of therapeutic intervention that is not compounded by the use of immunosuppressive drugs to prevent tissue rejection, while harnessing targeted gene-repair strategies, such as homologous recombination and zinc-finger nucleases, to repair genetic defects. These strategies provide the opportunity for generating an unlimited population of stem cells that can be differentiated into the desired cell type for studying disease mechanisms, for screening and developing drugs or for developing a suitable cell-replacement therapy. There have been considerable advances and successes to this end; however, selecting an appropriate disease target, directing the differentiation of iPS cells into phenotype-relevant cell populations and identifying disease-relevant phenotypes remain major hurdles. It is unclear whether iPS cells used for cell-replacement therapy would completely evade an immune response when returned to the patient, because a recent study has shown the immune rejection of teratomas formed from iPS cells, even in syngeneic mice[78]. Nevertheless, iPS cells provide a promising model with which to study disease mechanisms, discover new therapies and develop truly personalized treatments.

## Predictions for the evolution of the art

Few fields have enjoyed the remarkable upsurge in activity and excitement that followed the initial report of the reprogramming of somatic cells into iPS cells in 2006. Despite heady progress, crucial challenges must be met for the field to realize its full potential. There is as yet no consensus on the most consistent or optimal protocol for deriving the most reliable and, ultimately, the safest iPS cells. Increasing the reprogramming efficiency and effecting reprogramming without genetically modifying the cells are goals that have been achieved. Using more-uniform protocols and more-rigorous controls would facilitate experimental and potentially therapeutic consistency

| Disease | Molecular defect of donor cell | Cell type differentiated from iPS cells | Disease phenocopied in differentiated cells | Drug or functional tests |
|---|---|---|---|---|
| **Metabolic cont.** | | | | |
| Glycogen storage disease Ia (GSD1a) | Defect in glucose-6-phosphate gene | Hepatocyte-like cells (fetal) | Yes | No |
| Familial hypercholesterolaemia | Autosomal dominant mutation in *LDLR* | Hepatocyte-like cells (fetal) | Yes | No |
| Crigler–Najjar syndrome | Deletion in *UGT1A1* | Hepatocyte-like cells (fetal) | ND | No |
| Hereditary tyrosinaemia, type 1 | Mutation in *FAHD1* | Hepatocyte-like cells (fetal) | ND | No |
| Pompe disease | Knockout of *GAA* | Skeletal muscle cells | Yes | No |
| Progressive familial cholestasis | Multifactorial | Hepatocyte-like cells (fetal) | ND | No |
| Hurler syndrome (MPS IH) | Genetic defect in *IDUA* | Haematopoietic cells | No | No |
| **Cardiovascular** | | | | |
| LEOPARD syndrome | Heterozygous mutation in *PTPN11* | Cardiomyocytes | Yes | No |
| Type 1 long QT syndrome | Dominant mutation in *KCNQ1* | Cardiomyocytes | Yes | No |
| Type 2 long QT syndrome | Missense mutation in *KCNH2* | Cardiomyocytes | Yes | Yes |
| **Primary immunodeficiency** | | | | |
| SCID or leaky SCID | Mutation in *RAG1* | None | NA | No |
| Omenn syndrome (OS) | Mutation in *RAG1* | None | NA | No |
| Cartilage-hair hypoplasia (CHH) | Mutation in *RMRP* | None | NA | No |
| Herpes simplex encephalitis (HSE) | Mutation in *STAT1* or *TLR3* | Mature cell types of the central nervous system | No | No |
| **Other category** | | | | |
| Duchenne muscular dystrophy | Deletion in the dystrophin gene | None | NA | No |
| Becker muscular dystrophy | Unidentified mutation in dystrophin | None | NA | No |
| Dyskeratosis congenita (DC) | Deletion in *DKC1* | None | NA | No |
| Cystic fibrosis | Homozygous deletion in *CFTR* | None | NA | No |
| Friedreich's ataxia (FRDA) | Trinucleotide GAA repeat expansion in *FXN* | Sensory and peripheral neurons, and cardiomyocytes | Partially | No |
| Retinitis pigmentosa | Heterogeneity in causative genes and mutations: mutations in *RP9, RP1, PRPH2* or *RHO* | Retinal progenitors, photoreceptor precursors, retinal-pigment epithelial cells and rod photoreceptor cells | Yes | Yes |
| Recessive dystrophic epidermolysis bullosa (RDEB) | Mutation in *COL7A1* | Haematopoietic cells, and epidermis-like keratinocytes that differentiate into cells of all three germ layers *in vivo* | Partially | Yes |
| Scleroderma | Unknown | None | NA | No |
| Osteogenesis imperfecta | Mutation in *COL1A2* | None | NA | No |

An extended version of this table includes references and more information about drug and functional tests (Supplementary Table 1). *ABCD1*, ATP-binding cassette, sub-family D, member 1; *ADA*, adenine deaminase; *CFTR*, cystic fibrosis transmembrane conductance regulator; *COL1A2*, α2-chain of type I collagen; *COL7A1*, α1-chain of type VII collagen; *DKC1*, dyskerin; *FAA*, Fanconi's anaemia, complementation group A; *FAD2*, Fanconi's anaemia, complementation group D2; *FAHD1*, fumarylacetoacetate hydrolase; *FMR1*, fragile X mental retardation 1; *FXN*, frataxin; *GAA*, acid α-glucosidase; *GBA*, acid β-glucosidase; HbS, sickle haemoglobin; *HPRT1*, hypoxanthine phosphoribosyltransferase 1; *IDUA*, α-L-iduronidase; *JAK2*, Janus kinase 2; *KCNH2*, potassium voltage-gated channel, subfamily H (eag-related), member 2; *KCNQ1*, potassium voltage-gated channel, KQT-like subfamily, member 1; LDLR, low-density lipoprotein receptor; *LRRK2*, leucine-rich repeat kinase 2; *MECP2*, methyl CpG binding protein 2; NA, not applicable; *NAGLU*, α-N-acetylglucosaminidase; ND not determined; *PRPH2*, peripherin 2; *PTPN11*, protein tyrosine phosphatase, non-receptor type 11; *RAG1*, recombination activating gene 1; *RHO*, rhodopsin; *RMRP*, RNA component of mitochondrial-RNA-processing endoribonuclease; *RP*, retinitis pigmentosa; SCID, severe combined immunodeficiency; *SMN1*, survival of motor neuron 1; *SNCA*, α-synuclein; SOD1, superoxide dismutase 1; *STAT1*, signal transducer and activator of transcription 1; *TLR3*, Toll-like receptor 3; *UGT1A1*, UDP glucuronosyltransferase 1 family, polypeptide A1.

between laboratories and would yield standardized cell lines that could be used with confidence in both basic and applied studies.

Barring that, researchers must agree on standards of molecular analysis to ensure that the reprogrammed cells that most closely approximate the generic state of the naive genome can be identified. Because iPS cells are subject to the same type of culture adaptations that affect karyotypic integrity as human ES cells[79], it is important to define protocols that minimize the time in culture. In addition, cell lines used in clinical applications will need to be evaluated frequently for aberrant culture-induced changes at all stages: from the somatic cells to the reprogrammed and differentiated cells[80]. Understanding the genomic alterations that take place during the reprogramming, culture and differentiation of iPS cells will be crucial for designing experiments and ensuring that the derived cells are functional, pure and appropriate for use in research and therapy. Minimizing any aberrations is important, but as long as researchers understand that aberrations will arise — and can describe and control their effects — even imperfect cells can be used, and preferential differentiation can be taken advantage of whenever possible. Characteristics of iPS cells that were initially perceived as flaws, including varying differentiation propensities, might prove useful in clinical settings to generate cell types that have been difficult to obtain thus far.

Generating more stringent markers of pluripotency and assays to distinguish the abilities of a given iPS cell line are key priorities. Building on the progress that has already been made using ES cells[81], researchers must continue to improve the understanding of directed differentiation and to develop new protocols. With refined differentiation protocols, researchers will be able to investigate the pathophysiological basis of genetic diseases and carry out drug screening on affected cell types. These protocols will bring the field a step closer to patient-matched cells and tissues for clinical transplantation, a long-standing ambition of the stem-cell field that might be its ultimate measure of success. ■

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–676 (2006).
**This breakthrough paper describes the derivation of iPS cells directly from mouse somatic cells through the ectopic co-expression of reprogramming transcription factors, providing an alternative source of pluripotent cells for research.**
2. Dreisch, H. Entwicklungsmechanische Studien I. Der Wert der ersten beiden Furchungszellen in der Echinodermenentwickelung. Experimentelle Erzeugung von Teil und Doppelbildungen. *Z. Wiss. Zool.* **53,** 160–183 (1891).
3. Dewey, M. J., Martin, D. W. Jr, Martin, G. R. & Mintz, B. Mosaic mice with teratocarcinoma-derived mutant cells deficient in hypoxanthine phosphoribosyltransferase. *Proc. Natl Acad. Sci. USA* **74,** 5564–5568 (1977).
4. Gardner, R. L. Mouse chimeras obtained by the injection of cells into the blastocyst. *Nature* **220,** 596–597 (1968).
5. Brinster, R. L. The effect of cells transferred into the mouse blastocyst on subsequent development. *J. Exp. Med.* **140,** 1049–1056 (1974).
6. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292,** 154–156 (1981).
7. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78,** 7634–7638 (1981).
8. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282,** 1145–1147 (1998).
9. Vogel, G. & Holden, C. Ethics questions add to concerns about NIH lines. *Science* **321,** 756–757 (2008).
10. Mosher, J. T. *et al.* Lack of population diversity in commonly used human embryonic stem-cell lines. *N. Engl. J. Med.* **362,** 183–185 (2010).
11. Tabar, V. *et al.* Therapeutic cloning in individual parkinsonian mice. *Nature Med.* **14,** 379–381 (2008).
12. Noggle, S. *et al.* Human oocytes reprogram somatic cells to a pluripotent state. *Nature* **478,** 70–75 (2011).
13. Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451,** 141–146 (2008).
14. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–872 (2007).
15. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318,** 1917–1920 (2007).
16. Stadtfeld, M., Brennand, K. & Hochedlinger, K. Reprogramming of pancreatic β cells into induced pluripotent stem cells. *Curr. Biol.* **18,** 890–894 (2008).
17. Eminli, S., Utikal, J., Arnold, K., Jaenisch, R. & Hochedlinger, K. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous *Sox2* expression. *Stem Cells* **26,** 2467–2474 (2008).
18. Kim, J. B. *et al.* Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* **454,** 646–650 (2008).
19. Hanna, J. *et al.* Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133,** 250–264 (2008).
20. Aoi, T. *et al.* Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321,** 699–702 (2008).
21. Utikal, J., Maherali, N., Kulalert, W. & Hochedlinger, K. Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. *J. Cell Sci.* **122,** 3502–3510 (2009).
22. Sun, N. *et al.* Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl Acad. Sci. USA* **106,** 15720–15725 (2009).
23. Maherali, N. *et al.* A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3,** 340–345 (2008).
24. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448,** 196–199 (2007).
25. Wray, J., Kalkan, T. & Smith, A. G. The ground state of pluripotency. *Biochem. Soc. Trans.* **38,** 1027–1032 (2010).
26. Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143,** 508–525 (2010).
27. Maherali, N. & Hochedlinger, K. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* **3,** 595–605 (2008).
28. Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* **2,** 230–240 (2008).
**This study enumerated the molecular markers and functional criteria for defining reprogrammed cell populations.**
29. Chan, E. M. *et al.* Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nature Biotechnol.* **27,** 1033–1037 (2009).
30. Payer, B., Lee, J. T. & Namekawa, S. H. X-inactivation and X-reactivation: epigenetic hallmarks of mammalian reproduction and pluripotent stem cells. *Hum. Genet.* **130,** 265–280 (2011).
31. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1,** 55–70 (2007).
32. Stadtfeld, M. *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465,** 175–181 (2010).
33. Boland, M. J. *et al.* Adult mice generated from induced pluripotent stem cells. *Nature* **461,** 91–94 (2009).
34. Zhao, X. Y. *et al.* iPS cells produce viable mice through tetraploid complementation. *Nature* **461,** 86–90 (2009).
35. Eakin, G. S., Hadjantonakis, A. K., Papaioannou, V. E. & Behringer, R. R. Developmental potential and behavior of tetraploid cells in the mouse embryo. *Dev. Biol.* **288,** 150–159 (2005).
36. Eggan, K. & Jaenisch, R. Differentiation of $F_1$ embryonic stem cells into viable male and female mice by tetraploid embryo complementation. *Methods Enzymol.* **365,** 25–39 (2003).
37. Lensch, M. W., Schlaeger, T. M., Zon, L. I. & Daley, G. Q. Teratoma formation assays with human embryonic stem cells: a rationale for one type of human-animal chimera. *Cell Stem Cell* **1,** 253–258 (2007).
38. Park, I. H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134,** 877–886 (2008).
**This study derived iPS cells from patients with a range of diseases, demonstrating the applications of iPS cells for disease modelling, pathogenesis studies and drug development.**
39. Daley, G. Q. *et al.* Broader implications of defining standards for the pluripotency of iPSCs. *Cell Stem Cell* **4,** 200–202 (2009).
40. Miura, K. *et al.* Variation in the safety of induced pluripotent stem cell lines. *Nature Biotechnol.* **27,** 743–745 (2009).
41. Feng, Q. *et al.* Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells* **28,** 704–712 (2010).
42. Hu, B. Y. *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl Acad. Sci. USA* **107,** 4335–4340 (2010).
43. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467,** 285–290 (2010).
**This report describes an exhaustive comparison of pluripotent stem cells derived from mouse embryos or reprogrammed through nuclear transfer or transcription factor co-expression, by using several *in vitro* differentiation assays and methylation analysis; it revealed that iPS cells manifest molecular and behavioural features of the donor tissue of origin, indicating a 'memory' of the somatic tissue.**
44. Hu, Q., Friedrich, A. M., Johnson, L. V. & Clegg, D. O. Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. *Stem Cells* **28,** 1981–1991 (2010).
45. Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet Beta cells. *Cell Stem Cell* **9,** 17–23 (2011).
46. Urbach, A., Bar-Nur, O., Daley, G. Q. & Benvenisty, N. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell* **6,** 407–411 (2010).
47. Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell **144,** 439–452 (2011).
**This study generated genome-wide reference maps of DNA methylation and gene expression, together with the differentiation potential of each cell line, providing a resource for assessing the similarity of ES cells and iPS cells, as well as for predicting the differentiation efficiency of a particular cell line and creating a scorecard for the comprehensive characterization of any pluripotent cell line.**
48. Pick, M. *et al.* Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. *Stem Cells* **27,** 2686–2690 (2009).
49. Ben-David, U., Mayshar, Y. & Benvenisty, N. Large-scale analysis reveals acquisition of lineage-specific chromosomal aberrations in human adult stem cells. *Cell Stem Cell* **9,** 97–102 (2011).
50. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471,** 58–62 (2011).
**This study showed that significantly more CNVs were present in early-passage human iPS cells than in intermediate-passage human iPS cells, fibroblasts or human ES cells; it also provided evidence that CNVs conferred a selective disadvantage.**
51. Laurent, L. C. *et al.* Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* **8,** 106–118 (2011).
52. Ghosh, Z. *et al.* Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE* **5,** e8975 (2010).
53. Wernig, M. *et al.* A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nature Biotechnol.* **26,** 916–924 (2008).
54. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454,** 49–55 (2008).
55. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471,** 68–73 (2011).
**This study analysed the methylomes of human iPS cells, ES cells, somatic cells, and differentiated iPS and ES cells, and revealed megabase-scale DMRs in iPS cells, indicating incomplete reprogramming of these cells.**
56. Shen, Y. *et al.* X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proc. Natl Acad. Sci. USA* **105,** 4709–4714 (2008).
57. Tchieu, J. *et al.* Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell* **7,** 329–342 (2010).
58. Marchetto, M. C. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143,** 527–539 (2010).
59. Pomp, O. *et al.* Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. *Cell Stem Cell* **9,** 156–165 (2011).
60. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125,** 315–326 (2006).

61. Newman, A. M. & Cooper, J. B. Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* **7,** 258–262 (2010).
62. Humpherys, D. *et al.* Epigenetic instability in ES cells and cloned mice. *Science* **293,** 95–97 (2001).
63. Soldner, F. *et al.* Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* **136,** 964–977 (2009).
64. Carey, B. W. *et al.* Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* **9,** 588–598 (2011).
65. Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biol.* **13,** 541–549 (2011).
66. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnol.* **28,** 848–855 (2010).
67. Martinez, Y. *et al.* Cellular diversity within embryonic stem cells: pluripotent clonal sublines show distinct differentiation potential. *J. Cell. Mol. Med.* http://dx.doi.org/10.1111/j.1582-4934.2011.01334.x (in the press).
68. Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature Biotechnol.* **26,** 313–315 (2008).
69. Muller, F. J. *et al.* A bioinformatic assay for pluripotency in human cells. *Nature Methods* **8,** 315–317 (2011).
70. Lee, G. *et al.* Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461,** 402–406 (2009).
71. Moretti, A. *et al.* Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N. Engl. J. Med.* **363,** 1397–1409 (2010).
72. Itzhaki, I. *et al.* Modelling the long QT syndrome with induced pluripotent stem cells. *Nature* **471,** 225–229 (2011).
    **This study generated iPS cells from patients with long QT syndrome and stimulated them to differentiate into cardiomyocytes that paralleled the disease phenotype *in vitro*, and these cells were then used to evaluate the potency of existing and new therapeutic agents.**
73. Agarwal, S. *et al.* Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. *Nature* **464,** 292–296 (2010).
74. Batista, L. F. *et al.* Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. *Nature* **474,** 399–402 (2011).
75. Agarwal, S. & Daley, G. Q. Telomere dynamics in dyskeratosis congenita: the long and the short of iPS. *Cell Res.* **21,** 1157–1160 (2011).
76. Hanna, J. *et al.* Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **318,** 1920–1923 (2007).
77. Wernig, M. *et al.* Neurons derived from reprogrammed fibroblasts functionally integrate into the fetal brain and improve symptoms of rats with Parkinson's disease. *Proc. Natl Acad. Sci. USA* **105,** 5856–5861 (2008).
78. Zhao, T., Zhang, Z. N., Rong, Z. & Xu, Y. Immunogenicity of induced pluripotent stem cells. *Nature* **474,** 212–215 (2011).
79. Harrison, N. J., Baker, D. & Andrews, P. W. Culture adaptation of embryonic stem cells echoes germ cell malignancy. *Int. J. Androl.* **30,** 275–281 (2007).
80. Mayshar, Y. *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7,** 521–531 (2010).
81. Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132,** 661–680 (2008).
82. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105,** 2883–2888 (2008).
83. Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Biotechnol.* **26,** 1269–1275 (2008).
84. Sommer, C. A. *et al.* Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27,** 543–549 (2009).
85. Anokye-Danso, F. *et al.* Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8,** 376–388 (2011).
86. Woltjen, K. *et al. piggyBac* transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458,** 766–770 (2009).
87. Somers, A. *et al.* Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. *Stem Cells* **28,** 1728–1740 (2010).
88. Zhou, W. & Freed, C. R. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells* **27,** 2667–2674 (2009).
89. Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. Induced pluripotent stem cells generated without viral integration. *Science* **322,** 945–949 (2008).
90. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322,** 949–953 (2008).
91. Si-Tayeb, K. *et al.* Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev. Biol.* **10,** 81 (2010).
92. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn Acad.* **85,** 348–362 (2009).
93. Kim, D. *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4,** 472–476 (2009).
94. Zhou, H. *et al.* Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4,** 381–384 (2009).
95. Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7,** 618–630 (2010).
96. Miyoshi, N. *et al.* Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* **8,** 633–638 (2011).
97. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl Acad. Sci. USA* **107,** 9222–9227 (2010).
98. Buecker, C. *et al.* A murine ESC-like state facilitates transgenesis and homologous recombination in human pluripotent stem cells. *Cell Stem Cell* **6,** 535–546 (2010).
99. Li, W. *et al.* Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. *Cell Stem Cell* **4,** 16–19 (2009).
100. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4,** 487–492 (2009).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** G.Q.D. is supported by grants from the National Institutes of Health (RO1-DK70055, RO1-DK59279 and UO1-HL100001, as well as special funds from the American Recovery and Reinvestment Act of 2009 (RC2-HL102815), the Roche Foundation for Anemia Research, Alex's Lemonade Stand Foundation and the Harvard Stem Cell Institute. G.Q.D. is an affiliate member of the Broad Institute, a recipient of Clinical Scientist Awards in Translational Research from the Burroughs Wellcome Fund and the Leukemia and Lymphoma Society, and an investigator of the Howard Hughes Medical Institute and the Manton Center for Orphan Disease Research. We gratefully acknowledge A. de Los Angeles for providing the images for Fig. 1a, d, f; T. Onder, for Fig. 1b; and A. Cherry, for Fig. 1c, e.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The author declares competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to G.D. (george.daley@childrens.harvard.edu).

REVIEW

doi:10.1038/nature10762

# Clonal evolution in cancer

Mel Greaves[1] & Carlo C. Maley[2]

**Cancers evolve by a reiterative process of clonal expansion, genetic diversification and clonal selection within the adaptive landscapes of tissue ecosystems. The dynamics are complex, with highly variable patterns of genetic diversity and resulting clonal architecture. Therapeutic intervention may destroy cancer clones and erode their habitats, but it can also inadvertently provide a potent selective pressure for the expansion of resistant variants. The inherently Darwinian character of cancer is the primary reason for this therapeutic failure, but it may also hold the key to more effective control.**

Cancer is a major cause of death throughout the world and, despite an extraordinary amount of effort and money spent, the eradication or control of advanced disease has not been achieved[1]. Although we have a much greater understanding of cancer biology and genetics[2], translation into clinical practice needs to allow for the cellular complexity of the disease and its dynamic, evolutionary characteristics. These features provide both barriers to, and opportunities for, successful treatment.

In 1976, Peter Nowell[3] published a landmark perspective on cancer as an evolutionary process that is driven by stepwise, somatic-cell mutations with sequential, subclonal selection. This is a parallel to Darwinian natural selection, with cancer clones as the equivalent of asexually reproducing, unicellular quasi-species. Modern cancer biology and genomics have validated cancer as a complex, Darwinian, adaptive system[4,5] (Box 1 and Supplementary Information).

Cancer-clone evolution takes place within tissue ecosystem habitats. These habitats have evolved over a billion years to optimize multicellular function but restrain clonal expansion of renegade cells. However, the resilience of multicellular and long-lived animals depends on the phenotypic properties that, if not tightly regulated, drive or sustain malignancy: that is, cellular self-renewal and stabilization of telomeres, which allow extensive proliferation, angiogenesis, cell migration and invasion[6].

The long time period usually required for cancer symptoms to emerge and the complexity of the resultant mutations is, in part, a reflection of the sequential and random searches for phenotypic solutions to constraints from the micro-environment. The evolutionary progression of cancer is usually stalled or aborted, as shown by the high frequency of clinically covert premalignant lesions[7–9]. Cancer-suppressive mechanisms relegate most cancers to old age, when they have little effect on the reproductive fitness of their hosts.

Limited resources, environment architecture and other constraints of the micro-environment limit the size of solid tumours at every stage of their progression. Even advanced malignancies can show Gompertzian growth[10] — the cancer cell doubling time (around 1–2 days) is orders of magnitude faster than tumour doubling time (around 60–200 days)[10] — implying that the vast majority of cancer cells either die before they can divide[11] or are kept from dividing by the tumour micro-environment. Thus, natural selection in tumours, in the same way as selection in organisms, takes place through competition for space and resources.

Oncologists change cancer-clone dynamics by introducing a potent source of artificial selection in the form of drugs or radiation, but evolutionary principles still apply. Usually, treatment will result in massive cell death, which provides a selective pressure for the proliferation of variant cells that resist treatment (the mechanisms for this are discussed later). Furthermore, many cancer therapeutics are genotoxic; cells surviving treatment, which could then go on to regenerate the cancer, may have mutated further, resulting in cells with improved fitness and malignant potential.

The tools of and insights from evolutionary biology and ecology can therefore be applied to the dynamics of cancer before and after treatment to explain the modest returns from cancer therapy. We show that cancer is an inherently evolutionary process and suggest alternative strategies for effective control.

## Mutational drivers and clonal dynamics

The basic principle of a Darwinian evolutionary system is the purposeless genetic variation of reproductive individuals who are united by common descent, together with natural selection of the fittest variants. Cancer is a clear example of such a system. Most mutational processes have a bias at the DNA sequence level. The particular mutational spectra in a cancer cell can be a reflection of error-prone repair processes or associated with a genotoxic exposure (for example, cigarette carcinogens, ultraviolet light and chemotherapeutic drugs[2]). The patterns of genetic instability (chromosomal or microsatellite) in cancer cells may reflect exposure to, and the selective pressure exerted by, some classes of chemical carcinogens[2]. Nevertheless, for the functions encoded in genes, mutagenic processes are essentially blind or non-purposeful (with the exception of intrinsic mutagenic or recombinatorial enzymes preferentially targeting lymphoid immunoglobin or T-cell receptor genes[12]). The recurrent, mutation-endowed fitness traits in cancer reflect the potent impact clonal selection can have.

Clones evolve through the interaction of selectively advantageous 'driver' lesions, selectively neutral 'passenger' lesions and deleterious lesions (a 'hitchhiker' mutation in evolutionary biology is equivalent to a passenger mutation in cancer biology). In addition, 'mutator' lesions increase the rate of other genetic changes[13,14], and micro-environmental[15] changes alter the fitness effects of those lesions. The identification of driver lesions is supported by the independent observation that these lesions occur more frequently in multiple neoplasms than would be expected in the normal background mutation rate, that they are associated with clonal expansions[16,17] and from the type of mutation seen (missense, nonsense, frameshift, splice site, phosphorylation sites and double deletions)[18–20], particularly if the gene involved has a known role in cellular processes relevant to oncogenesis. The evidence gained from genetic studies

---

[1]Division of Molecular Pathology, The Institute of Cancer Research, Brookes Lawley Building, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK. [2]Center for Evolution and Cancer, Helen Diller Family Comprehensive Cancer Center, Department of Surgery, University of California, 2340 Sutter Street PO Box 1351, San Francisco, California 94115, USA.

of human tumours should be corroborated with functional tests and animal models. Passenger lesion status can also be ambiguous or context-dependent: for example, cases of monoallelic loss that only impact on function when the second allele of the same gene is lost, mutations that only cause a phenotypic effect when another gene locus also mutates, or cases in which the mutants are functionally relevant only in the context of therapeutic responses involving that gene.

Only a few studies have attempted to quantify the selective advantage provided by driver mutations. Bozic *et al.*[21] (using a non-spatial population genetics model of sequential, exponential clonal expansion) derived a formula for the proportion of expected neutral passenger mutations versus the proportion of selectively advantageous driver mutations as a function of the selective advantage of the driver mutations. By fitting this equation to glioblastoma and pancreatic cancer resequencing data, the authors estimated that driver mutations gave an average fitness advantage of only 0.4% (ref. 21). To measure the mutant clone selective advantage directly would require longitudinal samples of a neoplasm and estimation of the clone sizes at each time point.

The dynamics of somatic evolution depend on the interaction of mutation rate and clonal expansion. Mutation rate varies substantially between different genomic regions[22] and between different types of abnormality (for example, single-base sequence changes versus balanced chromosomal rearrangements and gene fusions), and mutation rates will increase by the instigation of genetic instability[23–25]. The rate of epigenetic change has been estimated to be orders of magnitude higher than that of genetic change[26], and could be a major determinant of clonal evolution. Natural selection affects epigenetic variation within neoplasms[27], because epigenetic changes are inherited at cell division and can affect cell phenotypes. Evolutionary biology tools to address many of these mutation rate complexities exist (see Supplementary Information), but these remain under used in cancer biology[28]. The traditional model of clonal evolution suggests that a series of clonal expansions grows to dominate the neoplasm ('selective sweeps')[16,21,29], but this can occur only if the time to the next driver mutation is longer than the time required for a clone to sweep through the neoplasm. In addition, if the second mutation occurs in a competitor clone, the expansion of both clones is restrained by mutual competition (known as clonal interference)[30]. Given the large population size and high mutation rate typical of neoplasms, clonal competition is probably common[31,32]. This issue is best addressed by serial sampling, and limited data suggest that parallel clonal expansions occur before subclones begin to dominate in early cancer development[33–35]. Initial evidence indicates that large clonal expansions after cell transformation are rare[26]. Direct evidence, from serial sampling of oncogenic mutations in advanced disease[36], metastasis[37] or post-chemotherapy relapses (see Supplementary Information), indicates selective sweeps originate from pre-existing genetic variants or subclones.

### Punctuated equilibrium versus gradualism

The argument of gradualism versus punctuated equilibrium[38] (a longstanding debate in species evolution) has recently emerged in the consideration of the clonal evolution of neoplasms. It is unknown whether malignant clones, with their markedly altered genomes, evolve gradually through a sequence of genetic alterations and clonal expansions; accumulate many lesions over time in a rare, undetected subclone that finally appears in a clonal expansion; or have a few, large-scale punctuated changes, possibly prompted by an acute insult or a single, catastrophic mitotic event that generates multiple lesions across the genome (or on a single chromosome, known as chromothripsis)[39]. Evidence of tens of non-synonymous mutations in cancers was interpreted under the assumption that they were generated by tens of clonal expansions[29]. Reconstruction of genealogies of neoplastic clones, based on genetic heterogeneity within neoplasms, suggests that clones with ancestral genomes are not driven to extinction by later clonal expansions[31–33], which allows the history of a neoplasm to be revealed. Breast cancer data[32] have shown that clones with intermediate genotypes are difficult to detect; each clone generates a cloud of genetic neutral or non-viable subclone variants around it. A study of B-cell chronic lymphocytic leukaemia[40] suggests that intermediate clones can be detected, but at a frequency of <0.001, which was below the detection threshold of the breast cancer study[32]. Intermediate clones may be rare because they have had limited potential to expand or because they were once common but were outcompeted by more recent clones.

The frequency of premalignant clonal lesions (or carcinoma *in situ*) substantially exceeds clinical cancer rates[7–9]. This, as well as cancer dormancy[41] and genetic reconstitution of clonal histories[37], indicates that cancer clones have long periods of stasis. However, cancer-clone evolution probably passes a point of no return, possibly at the metastatic growth stage. If unlimited proliferative capacity is guaranteed by telomere stabilization[25], then clonal expansion is stopped only when the size threatens the life of the patient. When provided (albeit rarely) with the routes for dissemination and immunoselection, cancer cells can have a parasite-like immortality and can re-establish themselves in other individuals[6,42,43].

**BOX 1**

## Cancer as a complex system

- Cancers exist in a variety of taxonomic quasi-classes, genera, species, characterized by divergent cells of origin and mutational spectra. Each cancer is unique.
- Cancers evolve over a variable time frame (anywhere from 1 to 50 years), and the clonal structure, genotype and phenotype can shift over time in each patient. Each cancer is, in effect, multiple different (subclonal) cancers that occupy overlapping or distinct tissue habitats.
- The number of mutations in a cancer can vary from a handful (10–20) to (the more usual) hundreds or thousands. The great majority are passengers, and a modest, but undefined, number are functionally relevant drivers. The mutational processes are very diverse.
- Cancers acquire, through mutational and epigenetic changes, a variety of phenotypic traits that compound to allow territorial expansion, by proliferative self-renewal, migration and invasion — properties that are cruical to normal developmental, physiological and repair processes.
- Advanced, disseminated or very malignant cancers seem to be almost uniquely competent to evade therapy.
- Most, if not all, of this complexity can be explained by classical evolutionary principles.

## The cancer ecosystem

Tissue ecosystems provide the venue and determinants for fitness selection (the adaptive landscape[44]). Tissue micro-environments are complex, dynamic states with multiple components that can influence cancer-clone evolution (Fig. 1). For example, transforming growth factor-β is a cancer-ecosystem regulatory molecule[45]. Other cellular and cytokine components of inflammatory lesions are potent and common modulators of the cancer-cell ecosystem[25].

The interaction between cancer cells and their tissue habitats is reciprocal. Cancer cells can remodel tissue micro-environments and specialized niches to their competitive advantage[46]. Cancer-clone expansion is controlled by architectural constraints or barriers, such as sequestration of stem cells into crypts in the gastrointestinal tract[47]

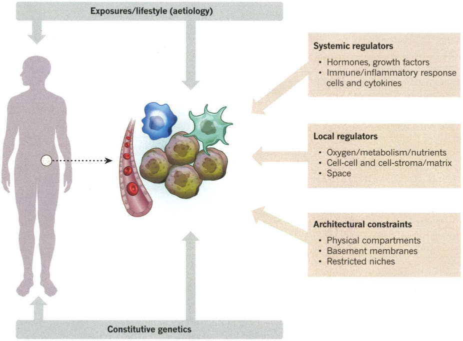

Exposures/lifestyle (aetiology)
Systemic regulators
• Hormones, growth factors
• Immune/inflammatory response cells and cytokines
Local regulators
• Oxygen/metabolism/nutrients
• Cell-cell and cell-stroma/matrix
• Space
Architectural constraints
• Physical compartments
• Basement membranes
• Restricted niches
Constitutive genetics

**Figure 1 | The complexity of tissue ecosystems.** Exposure, the constitutive genetics of the host cells, systemic regulators, local regulators and architectural constraints all impinge on the evolution of somatic cells.

and the need for external signals for proliferation and cell survival. However, some micro-environmental components can promote neoplastic cells; for example, infiltrating macrophages and neovascularization, in response to anoxia, can support neoplastic cell survival and proliferation. Mathematical modelling shows cancer-clone evolutionary selection for more robust or malignant phenotypes is less likely in more stable or homogeneous micro-environments[48]. Spatial heterogeneity of resources in the primary tumour selects for cell migration and emigration, which may explain why there is selection for metastasis[49]. Preclinical models have suggested that normalizing the resources across the primary tumour can suppress metastasis[50]. As clones and subclones expand, migrant cells invade new habitats within and between tissues, in which they experience new selective pressures that can cause further cancer-cell diversity. This malignant feature, and its associated morbidity, characterizes end-stage cancer.

Cancer-cell habitats are not closed systems. The tissue ecosystem, in addition to regulation by systemic factors (such as nutrients and hormones) or invasion by inflammatory or endothelial cells, is modified by external factors. As well as the tissue site, the ecosystem for each cancer includes environmental, lifestyle and associated aetiological exposure of the patient. Genotoxic exposure (such as, cigarette carcinogens or ultraviolet light), infection, and long-term dietary and exercise habits that affect calorie, hormone or inflammation levels can have a profound effect on the tissue micro-environments, as well as directly on cancer cells (Fig. 1). These factors are the aetiological link to the initiation or progression of cancer, and without such modulating exposure, the risk of cancer-clone initiation and evolution would be reduced.

Cancer-tissue ecosystems can be radically altered after chemotherapy or radiotherapy. Most cancer cells may be decimated, but the remodelled landscape creates new selective pressures, resources and opportunities that may allow pre-existing variant cancer cells that survived treatment to emerge. Crucially, stroma or specialized habitat niches may protect cancer cells against the therapy[51].

## Cancer genomics and clonal architecture

Cancer-genome sequencing, facilitated by the introduction of second-generation whole-genome sequencing, has provided further insight into the complexity of the genetics and evolutionary biology of cancer cells[2]. In most cases, transformation and metastases are probably clonal[2] because they are derived from single cells; therefore, the identification of the mutations present in all of the cells of a tumour can help to reconstruct the genotype of the founder cell. These founder events limit the genetic and clonal complexity of tumours. We already had a long list of recurring driver mutations (with gain or loss of function) as a result of the fine mapping of chromosomal breaks, candidate gene sequencing and functional screening of bulk samples from tumours. However, the use of genomic screens has demonstrated the scale of cancer-genome complexity. Individual cancers can contain hundreds, or tens of thousands, of mutations and chromosomal alterations[2]. The great majority of these are assumed to be neutral mutations arising from genetic instability. Chromosomal instability (amplifications, deletions, translocations and other structural changes) is a common feature, but it is not clear whether there is an increased rate of simple base-pair mutations in cancer[2,21,23,52]. Evolutionarily neutral alterations are thought to register in the screens because they hitchhike on clonal expansions that are driven by selectively advantageous alterations or by drift. In addition, data have confirmed that each cancer in each patient has an individually unique genomic profile. It is possible that cancer cells need only a modest number of phenotypic traits to deal with all of the constraints and evolve into a fully malignant or metastatic tumour[25], but the genomics data suggest that this can be achieved by an almost infinite variety of evolutionary trajectories and with multiple different combinations of driver mutations[44].

Paradoxically, genome profiles underestimate complexity. So far, they have been mostly one-off snapshots from a single sample at a single diagnostic time point. We know that serial or parallel sampling using more conventional genetic analysis uncovers genetic diversity within a tumour. Whole-genome sequencing of paired primary tumours versus metastatic samples has so far been limited, but it has revealed that individual metastatic lesions are clonal in origin and genetically unique, yet have a clonal ancestry traceable to the primary tumour[2]. 'The genome' description is perhaps also misleading because genetic variants are identified in 5–50% of reads, which suggests subclonal distribution of most mutations[53], but the segregation pattern of mutations within subclones is lost when DNA is extracted from the total cell population. This is important if patient-specific genomic profiles are to provide a platform for selecting therapeutic targets. Arguably, subclonal genetic diversity is key to the success or failure of therapy. This is a considerable challenge, technically and bioinformatically, in cancer genomics and will require deep sequencing[40]

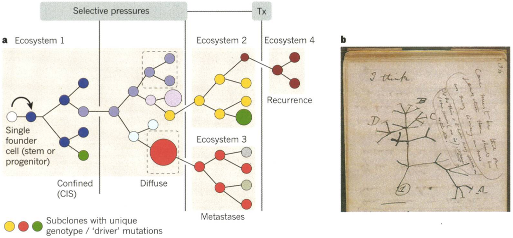


**Figure 2 | The branching architecture of evolution.** **a**, Cancer clones. Selective pressures allow some mutant subclones to expand while others become extinct or remain dormant. Vertical lines represents restraints or selective pressures. This is a representative pattern for common, solid cancers; as recognized by Nowell[3], leukaemic clones may expand over a shorter time frame (years versus decades), and be subject to fewer restraints and mutational events. Ecosystems 1–4 (boxes) represent the different tissue ecosystems or habitats. Smaller boxes within Ecosystem 1 represent localized habitats or niches. Each differently coloured circle represents a genetically distinct subclone. Metastatic subclones can branch off into different time points in the sequence from either minor or major clones in the primary tumour. Tx, therapy. CIS, carcinoma *in situ*. **b**, Darwin's branching evolutionary tree of speciation from his 1837 notebook.

and investigation of the genomes of single cells for patterns of segregation of mutations to understand the genetic diversity within neoplasms and how this changes in response to interventions.

## Subclonal segregation of mutations and clonal architecture

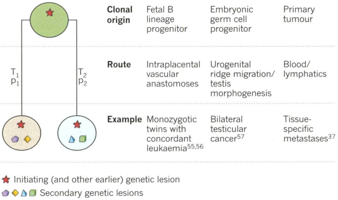
The classic model of clonal evolution suggests there is a sequential acquisition of mutations with concomitant, successive subclonal dominance or selective sweeps. Histopathological evidence of disease progression (adenoma, carcinoma and metastases) supports this model. At each stage of this evolution, individual cells and their progeny (subclones) compete for space and resources. Multiplexed, single-cell mutational analysis (ideally in serial samples) is the most appropriate way to examine clonal architecture. So far, there are only a few examples of this[10,32,33], but they have provided evidence of the complex pattern of subclonal segregation of mutations — consistent with Nowell's model. The large amount of data from tissue sections, small biopsies and, more recently, single-cell analysis[33] is evidence that the evolutionary trajectories are complex and branching, exactly as Nowell proposed and in parallel with Darwin's iconic evolutionary speciation tree (Fig. 2). Attempts to simplify this complex system into a linear sequence of mutational events on the basis of cross-sectional data have probably been misleading[54]. However, by comparing the mutational genomes of the subclones, it is possible to discover their evolutionary or ancestral relationships, as well as the order of events during the development of that neoplasm[32,33,37,53,54]. Clonal evolution from common ancestral cancer cells is demonstrated in identical twins with concordant acute leukaemia[55,56], in metastatic lesions[2,10] and, by inference, in some cases of bilateral testicular cancer[57] (Fig. 3). In this context, divergent cancer-clone genotypes and phenotypes correspond to allopatric speciation in separate natural habitats (for example, Darwin's finches on the Galapagos Islands[58]).

Profiles of subclones within a neoplasm can be used to determine 'molecular clocks' that can then be linked to time events in the history of the neoplasm. For example, DNA methylation changes and base-pair mutations have been used to infer clonal expansion dynamics[26] and the time between initiation, invasion and metastasis[17,37,52]. It is even possible to determine the relative timing of events during progression from a single sample, based on deep sequencing[59].

Subclones may be mixed together within the primary tissue[37,60], but given their single-cell origin and bifurcating pathways, it is not surprising that they can also occupy distinctive territories[35,37,61,62] (Fig. 4a). Cancer-clone evolution involves contemporaneous subclones with distinctive mutational and phenotypic profiles that may be territorially segregated, which has considerable practical implications for diagnosis, prognosis and targeted therapy based on biopsy sampling[63]. It remains unclear whether all subclonal diversification reflects the impact of driver mutations and selective advantage, or is also the result of genetic drift of selectively neutral mutations or even epigenetic alterations. The level of diversity within the subclonal structure can be measured[35,64,65] and has been shown to be a robust biomarker for predicting progression to malignancy in Barrett's oesophagus[65]. It is also associated with the tumour stage and subtype of breast cancer[64].

## Units of selection and cancer stem cells

Evolutionary theory suggests that natural selection operates in any system that has components with varying reproductive potential[4]. In



| | | | |
|---|---|---|---|
| **Clonal origin** | Fetal B lineage progenitor | Embryonic germ cell progenitor | Primary tumour |
| **Route** | Intraplacental vascular anastomoses | Urogenital ridge migration/ testis morphogenesis | Blood/ lymphatics |
| **Example** | Monozygotic twins with concordant leukaemia[55,56] | Bilateral testicular cancer[57] | Tissue-specific metastases[37] |

★ Initiating (and other earlier) genetic lesion
⬟ ◆ △ ▢ Secondary genetic lesions

**Figure 3 | Divergent (branching) clonal evolution of cancer with topographical separation.** In each example, a clonal (single cell) ancestry is indicated by a shared acquired mutation (for example, *ETV6–RUNX1* fusion for leukaemias and *KIT* mutation for testicular cancers). The time at which the two subclones evolve ($T_1$ and $T_2$) can be temporarily synchronous or develop several years apart[37,55–57]. The probabilities of subclones emerging as shown are independent and different ($p_1$ and $p_2$). In most cases (90% for monozygotic twins), only one twin develops overt leukaemia. The penetrance of bilateral testicular cancer having a common origin[57] is unknown.

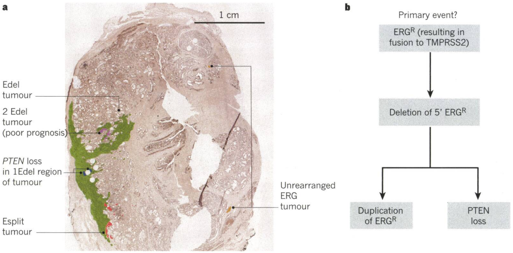


**Figure 4 | Topography of cancer subclones. a,** Tissue section of prostate to detect genetic events: *TMPRSS2–ERG* fusion (*ERG* via rearrangement (*ERG*$^R$)) and *PTEN* loss. **b,** The presumed sequence of clonal events.

the progression of cancer, or its resurgence after therapy, the primary unit of selection is the cell. This cell has to have extensive replicative potential, the so-called cancer stem cell (also known as the cancer-initiating or propagating cell) (Fig. 5).

The cancer stem-cell hypothesis was developed through transplantation experiments with leukaemic cells[66], and although it has been reported to be a general feature of all cancers[67], this idea is contentious. There has been no consensus on whether cancer stem cells are rare or high-frequency cells, or whether they have fixed, hierarchical or variable phenotypes, but considering the evolutionary progression in cancer, cells with extensive propagating activity are unlikely to be fixed entities[68,69]. Cancer stem cells are the cellular drivers of subclonal expansion and so probably vary in frequency and phenotypic features. The only feature they must have is the potential for extensive self-renewal (Fig. 5). Quantitative measures of stem-cell activity or self-renewal (through xenotransplantation or gene-expression signatures) can be used to predict the clinical outcome of several cancer types[70]. The cancer stem cell's ability to self-renew is made stronger by an aberrant genotype and, possibly, other, epigenetic, features. Several testable predictions can be made from this. First, cancer stem cells should evolve and change in genotype and phenotype as the cancer evolves before and after therapy. Some therapies may even provide a strong selection for cancer stem-cell survival and proliferation[71]. Second, as cancers progress, there should be selective pressure for the cells with the most extensive self-renewing capacity, but at the expense of cells with the ability to differentiate. This has been observed in chronic myeloid leukaemia (CML)[72] and mouse models[73,74]. A higher probability of symmetrical self-renewing proliferative cycles would be expected to result in an increased number and frequency of cancer stem cells. It is therefore of some consequence that loss of the *TP53* DNA damage checkpoint, which frequently correlates with cancer progression and clinical intransigence[75], seems to 'release' stem-cell-like transcriptional signatures[76] and leads to enhanced self-renewal in mammosphere culture systems[77]. The frequency of cancer stem cells could then increase from low to very high frequency as the disease progresses[78,79]. Third, for selection to operate through micro-environmental or therapeutic pressures, there should be contemporaneous genetic variation in cancer stem cells, which has been shown in leukaemias[33,80].

These considerations have significant clinical implications. Whatever the frequency and phenotype, if self-renewing cancer stem cells drive and sustain cancer-clone evolution, this suggests they are the repository of functionally relevant mutational events that drive clonal selection before and after therapy. This supports the view that cancer stem-cell restraint or elimination should be the aim of any therapy. However, if cancer stem cells are as genetically (and epigenetically) diverse as evolutionary considerations and initial experiments[33,80,81] indicate, this could be the reason for therapeutic failure. The adaptability of cancer stem cells provided by genetic diversity is added to by what seems to be their intrinsically lowered susceptibility to drugs and irradiation[82]. This may be because of the association with stromal cells[83] and the quiescence of cancer stem-cell subpopulations, as well as the properties of enhanced DNA repair and elevated expression of drug efflux pumps, which may be the evolved contingencies to protect normal stem cells.

Subclonal genetic heterogeneity is a common, if not universal, feature of cancers[84]. However, it cannot be assumed that all subclones are sustained by cancer stem cells; some could be evolutionary dead-ends generated by cells with only limited propagating potential. It is partly to accommodate this that the *in vivo* assay for cancer stem cells involves sequential transplants[66]. Ideally, the genomes of single cancer stem cells would be interrogated to investigate how they relate to subclones, but this is not currently possible. However, the genetic heterogeneity of cancer stem cells can be inferred by comparing subclonal diversity or clonal architecture before and after transplantation. Quadrant sections of glioblastoma have been shown to have divergent but related genotypes, but all sections contained cells that read-out in the *in vivo* (intracerebral) cancer stem-cell assay[85]. More definitive data come from comparing pre- and post-transplant subclonal genetic profiles that were investigated at the single cell level or by single nucleotide polymorphism arrays in B-cell precursor acute lymphoblastic leukaemia. Multiple subclones from each patient's diagnostic sample registered in the *in vivo* cancer stem-cell transplant assays, albeit with variable competitive potency[33,80,81]. We are still awaiting experimental confirmation that genetic diversity of cancer stem cells is a common feature of cancer, but, assuming that it is, this will have important therapeutic implications.

## A Darwinian bypass

Nowell[3] stated in his landmark article "more research should be directed towards understanding and controlling the evolutionary process in tumours before it reaches the late stage seen in clinical cancer". Although cancer therapy has had its successes, in reality very few advanced or metastatic malignancies can be effectively controlled or eradicated. Genetic variation in cancer stem cells, particularly if induced by genetic instability, provides the opportunity for cells to escape and the therapy to fail. Other, non-genetic, mechanisms of positive selection by therapy exist, including signalling plasticity (or oncogene bypass)[86], quiescence[87] and epigenetic changes[88]; however, many of these depend on heritable, and thus selectable, epigenetic variation. Great expectation has been placed on the audit of cancer genomes that, by identifying recurrent and "druggable" mutations, would herald a new phase of highly specific or targeted small-molecule inhibitors and personalized medicine[89]. Oncogene addiction may be the Achilles heel of cancer in this respect[90]. The success of imatinib and the derivative non-receptor tyrosine (ABL1) kinase inhibitors in CML[90] was very encouraging, but CML is not a typical cancer. It is essentially a premalignant (albeit ultimately lethal) condition, probably driven by a single founder mutation (*BCR–ABL1* fusion), which provides a universal target for therapy. Even in the most favourable of circumstances, escape occurs either by quiescence (and coupled resistance) of cancer stem cells[91] or by mutation of the ABL1 kinase target. Once CML has evolved to an overt malignancy or blast crisis, with increased genetic complexity, ABL1 kinase-directed therapy is often ineffective.

Other small-molecule inhibitors directed at mutant products have produced encouraging results in patients with advanced disease, but the benefits are transitory and cancer clones re-emerge with resistant features. When the targets selected are non-founder mutations, even if they are dominant in the neoplasm, therapy can be predicted to select for subclones lacking the mutant target[70]. Alternatively, subclones can have additional mutations that allow a bypass of the signalling pathway of the drug target, such as the MET proto-oncogene (*MET*) amplification in *EGFR* mutant lung cancer treated with EGFR kinase inhibitors[92].

Supporters of targeted therapy and personalized medicine argue that a combination of drugs that target components of networked signalling and are tailored to the individual patient's cancer genome is the solution to this problem. In this regard, synthetic lethal strategies seem promising[93].

Self-renewing cancer cells are the ultimate target for therapy, so high-throughput screening for selective inhibitors is an encouraging development[71]. Ways to target the components of the self-renewing process itself (independent of specific mutant genotype) deserve exploration, especially if a distinction can be made from normal adult stem cells. In the case of CML, intrinsically resistant (and possibly quiescent) stem cells, have been targeted by combining selective kinase (ABL1) inhibitors with inhibitors of a histone deacetylase[94] or BCL6 (ref. 95). Ultimately, it may prove difficult to thwart the plasticity and adaptability of cancer cells (or cancer stem cells), which are an inherent evolutionary feature of advanced disease, and a 'Darwinian bypass' may be required, for which there are a number of possibilities. An implication of the evolutionary diversity of cancer is that prevention (smoking cessation, avoiding sunburn, prophylactic vaccines, and so on) makes a great deal of sense, as does early detection and intervention (that is, before genetic diversification and dissemination become extensive).

An alternative therapeutic strategy is to focus on the microenvironmental habitat using 'ecological therapy', which aims to change the essential habitat and dependency of the cancer cells[96]. For example, anti-angiogenesis can provide a potent restraint on cancer stem cells[97]. Other examples are the use of bisphosphonates to remodel bone in patients with prostate cancer, the use of aromatase inhibitors in patients with breast cancer, exploiting hypoxia, the use of inhibitors of inflammation or tumour-infiltrating macrophages, and blocking cancer stem-cell interactions with essential stromal or niche components[96,98].

Another alternative is to control the cancer, rather than eradicate it, thereby turning cancer into a chronic disease. Because the speed of evolution is proportional to the fitness differential between the cells, cytotoxic drugs are predicted to select rapidly for resistance[5]. It is thought they cause competitive release[99] by removing all of the competitors of resistant cells. In contrast, cytostatic drugs should delay progression and mortality longer than cytotoxic drugs because sensitive competitor cells remain in the tissue to occupy space and consume resources that would otherwise be used by the resistant clones. In addition, by suppressing cell division, cytostatic drugs also suppress the opportunities for new mutations. A study by Gatenby and colleagues[100] showed that by treating an aggressive ovarian cancer (OVCAR-3) xenograft tumour to maintain a stable size, rather than to eradicate it, host mice could be kept alive much longer. Moreover, the dose of carboplatin necessary to keep the tumour at a manageable size declined over time[100]. Researchers should now focus on what phenotypes can be selected for to make neoplasms less deadly and more clinically manageable.

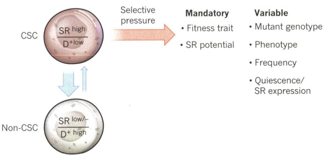


**Figure 5 | Selective pressure on cancer stem cells.** Selective pressures can include environmentally derived genotoxicity, natural or physiological restraints, cancer therapy, and so on. Mutation in progenitor cells can convert these cells back to a self-renewing population[72]. The small blue arrow represents a mutation; the large blue arrow represents differentiation: in both cases they represent a change in state. In addition to the mandatory trait of self-renewal, cancer stem cells (CSC), can exhibit any phenotypic feature that allows cells to continue to survive and proliferate in the face of a particular constraint. D$^+$, differentiation; SR, self-renewal.

The evolutionary theory of cancer has survived 35 years of empirical observation and testing, so today it could be considered a bona fide scientific theory. The basic components of somatic evolution are well understood, but the dynamics of somatic evolution remain unclear. Fortunately, there are evolutionary biology tools that may be applied to neoplasms to address many of the fundamental cancer biology questions, such as the order of events in progression, distinguishing driver from passenger mutations, and understanding and preventing therapeutic resistance. The dynamics of clonal diversification and selection are critical to understanding these issues. The challenge now is to use the clinical opportunities to address directly the evolutionary adaptability of neoplasms and design interventions to slow, direct or control cancer-cell evolution to delay or prevent mortality. ■

1. Jemal, A. *et al.* Cancer statistics, 2008. *CA Cancer J. Clin.* **58,** 71–96 (2008).
2. Stratton, M. R. Exploring the genomes of cancer cells: progress and promise. *Science* **331,** 1553–1558 (2011).
3. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194,** 23–28 (1976).
**The foundation paper that established the evolutionary theory of cancer.**
4. Merlo, L. M., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nature Rev. Cancer* **6,** 924–935 (2006).
5. Pepper, J., Scott Findlay, C., Kassen, R., Spencer, S. & Maley, C. Cancer research meets evolutionary biology. *Evol. Appl.* **2,** 62–70 (2009).
6. Greaves, M. *Cancer: The Evolutionary Legacy* (Oxford Univ. Press, 2000).
7. Sakr, W. A., Haas, G. P., Cassin, B. F., Pontes, J. E. & Crissman, J. D. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *J. Urol.* **150,** 379–385 (1993).
8. Mori, H. *et al.* Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc. Natl Acad. Sci. USA* **99,** 8242–8247 (2002).
9. Reid, B. J., Li, X., Galipeau, P. C. & Vaughan, T. L. Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nature Rev. Cancer* **10,** 87–101 (2010).
10. Klein, C. A. Parallel progression of primary tumours and metastases. *Nature Rev. Cancer* **9,** 302–312 (2009).
11. Malaise, E. P., Chavaudra, N. & Tubiana, M. The relationship between growth rate, labelling index and histological type of human solid tumours. *Eur. J. Cancer* **9,** 305–312 (1973).
12. Tsai, A. G. *et al.* Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell* **135,** 1130–1142 (2008).
13. Bardelli, A. *et al.* Carcinogen-specific induction of genetic instability. *Proc. Natl Acad. Sci. USA* **98,** 5770–5775 (2001).
14. Cahill, D. P., Kinzler, K. W., Vogelstein, B. & Lengauer, C. Genetic instability and Darwinian selection in tumors. *Trends Cell Biol.* **9,** M57–M60 (1999).

15. Barcellos-Hoff, M. H., Park, C. & Wright, E. G. Radiation and the microenvironment – tumorigenesis and therapy. *Nature Rev. Cancer* **5,** 867–875 (2005).
16. Maley, C. C. *et al.* Selectively advantageous mutations and hitchhikers in neoplasms: p16 lesions are selected in Barrett's esophagus. *Cancer Res.* **64,** 3414–3427 (2004).
17. Tao, Y. *et al.* Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. *Proc. Natl Acad. Sci. USA* **108,** 12042–12047 (2011).
18. Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463,** 893–898 (2010).
19. Youn, A. & Simon, R. Identifying cancer driver genes in tumor genome sequencing studies. *Bioinformatics* **27,** 175–181 (2011).
20. Greenman, C., Wooster, R., Futreal, P. A., Stratton, M. R. & Easton, D. F. Statistical analysis of pathogenicity of somatic mutations in cancer. *Genetics* **173,** 2187–2198 (2006).
21. Bozic, I. *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl Acad. Sci. USA* **107,** 18545–18550 (2010).
22. Schwartz, M., Zlotorynski, E. & Kerem, B. The molecular basis of common and rare fragile sites. *Cancer Lett.* **232,** 13–26 (2006).
23. Loeb, L. A. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nature Rev. Cancer* **11,** 450–457 (2011).
24. Weisenberger, D. J. *et al.* CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with *BRAF* mutation in colorectal cancer. *Nature Genet* **38,** 787–793 (2006).
25. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
**This paper consolidates the common phenotypes that evolve in neoplastic cells of all types.**
26. Siegmund, K. D., Marjoram, P., Woo, Y. J., Tavare, S. & Shibata, D. Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. *Proc. Natl Acad. Sci. USA* **106,** 4828–4833 (2009).
27. Varley, K. E., Mutch, D. G., Edmonston, T. B., Goodfellow, P. J. & Mitra, R. D. Intra-tumor heterogeneity of *MLH1* promoter methylation revealed by deep single molecule bisulfite sequencing. *Nucleic Acids Res.* **37,** 4603–4612 (2009).
28. Aktipis, C. A., Kwan, V. S. Y., Johnson, K. A., Neuberg, S. L. & Maley, C. C. Overlooking evolution: a systematic analysis of cancer relapse and therapeutic resistance research. *PLoS ONE* **6,** e261000 (2011).
29. Beerenwinkel, N. *et al.* Genetic progression and the waiting time to cancer. *PLoS Comput. Biol.* **3,** e225 (2007).
30. de Visser, J. A. & Rozen, D. E. Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. *Genetics* **172,** 2093–2100 (2006).
31. Leedham, S. J. *et al.* Individual crypt genetic heterogeneity and the origin of metaplastic glandular epithelium in human Barrett's oesophagus. *Gut* **57,** 1041–1048 (2008).
32. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472,** 90–94 (2011).
**Single-cell sequencing revealed the clonal structure of two breast cancers.**
33. Anderson, K. *et al.* Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469,** 356–361 (2011).
**Single-cell genetic analyses and xenografts revealed the clonal architecture within acute lymphoblastic leukaemia stem-cell populations and demonstrated repeated independent acquisition of copy number changes within the same neoplasm.**
34. Tsao, J. L. *et al.* Colorectal adenoma and cancer divergence. Evidence of multilineage progression. *Am. J. Pathol.* **154,** 1815–1824 (1999).
35. Maley, C. C. *et al.* Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genet.* **38,** 468–473 (2006).
36. Sidransky, D. *et al.* Clonal expansion of p53 mutant cells is associated with brain tumour progression. *Nature* **355,** 846–847 (1992).
37. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467,** 1114–1117 (2010).
38. Gould, S. J. & Eldredge, N. Punctuated equilibrium comes of age. *Nature* **366,** 223–227 (1993).
39. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144,** 27–40 (2011).
40. Campbell, P. J. *et al.* Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc. Natl Acad. Sci. USA* **105,** 13081–13086 (2008).
**Deep sequencing revealed rare (frequency <0.001) intermediate genotypes between the common clones in leukaemias (using immunoglobulin rearrangements as surrogate mutations).**
41. Aguirre-Ghiso, J. A. Models, mechanisms and clinical evidence for cancer dormancy. *Nature Rev. Cancer* **7,** 834–846 (2007).
42. Isoda, T. *et al.* Immunologically silent cancer clone transmission from mother to offspring. *Proc. Natl Acad. Sci. USA* **106,** 17882–17885 (2009).
43. Welsh, J. S. Contagious cancer. *Oncologist* **16,** 1–4 (2011).
44. Gatenby, R. A. & Gillies, R. J. A microenvironmental model of carcinogenesis. *Nature Rev. Cancer* **8,** 56–61 (2008).
45. Bierie, B. & Moses, H. L. Tumour microenvironment: TGFβ: the molecular Jekyll and Hyde of cancer. *Nature Rev. Cancer* **6,** 506–520 (2006).
46. Lathia, J. D., Heddleston, J. M., Venere, M. & Rich, J. N. Deadly teamwork: neural cancer stem cells and the tumor microenvironment. *Cell Stem Cell* **8,** 482–485 (2011).
47. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255,** 197–200 (1975).
**This paper identified natural selection as a driving force in carcinogenesis and identified tissue architecture as a cancer suppressor, and posited an immortal strand of DNA in tissue stem cells.**
48. Anderson, A. R., Weaver, A. M., Cummings, P. T. & Quaranta, V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **127,** 905–915 (2006).
49. Chen, J., Sprouffske, K., Huang, Q. & Maley, C. C. Solving the puzzle of metastasis: the evolution of cell migration in neoplasms. *PLoS ONE* **6,** e17933 (2011).
50. Mazzone, M. *et al.* Heterozygous deficiency of *PHD2* restores tumor oxygenation and inhibits metastasis via endothelial normalization. *Cell* **136,** 839–851 (2009).
51. Gilbert, L. A. & Hemann, M. T. DNA damage-mediated induction of a chemoresistant niche. *Cell* **143,** 355–366 (2010).
52. Jones, S. *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl Acad. Sci. USA* **105,** 4283–4288 (2008).
53. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464,** 999–1005 (2010).
54. Sprouffske, K., Pepper, J. W. & Maley, C. C. Accurate reconstruction of the temporal order of mutations in neoplastic progression. *Cancer Prev. Res.* **4,** 1135–1144 (2011).
55. Greaves, M. F., Maia, A. T., Wiemels, J. L. & Ford, A. M. Leukemia in twins: lessons in natural history. *Blood* **102,** 2321–2333 (2003).
56. Bateman, C. M. *et al.* Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115,** 3553–3558 (2010).
57. Oosterhuis, J. W. & Looijenga, L. H. Testicular germ-cell tumours in a broader perspective. *Nature Rev. Cancer* **5,** 210–222 (2005).
58. Grant, P. R. & Grant, B. R. *How and Why Species Multiply* (Princeton Univ. Press, 2008).
59. Durinck, S. *et al.* Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov.* **1,** 137–143 (2011).
60. Gonzalez-Garcia, I., Sole, R. V. & Costa, J. Metapopulation dynamics and spatial heterogeneity in cancer. *Proc. Natl Acad. Sci. USA* **99,** 13085–13089 (2002).
61. Clark, J. *et al.* Complex patterns of *ETS* gene alteration arise during cancer development in the human prostate. *Oncogene* **27,** 1993–2003 (2008).
62. Navin, N. *et al.* Inferring tumor progression from genomic heterogeneity. *Genome Res.* **20,** 68–80 (2010).
63. Allred, D. C. *et al.* Ductal carcinoma *in situ* and the emergence of diversity during breast cancer evolution. *Clin. Cancer Res.* **14,** 370–378 (2008).
64. Park, S. Y., Gonen, M., Kim, H. J., Michor, F. & Polyak, K. Cellular and genetic diversity in the progression of *in situ* human breast carcinomas to an invasive phenotype. *J. Clin. Invest.* **120,** 636–644 (2010).
65. Merlo, L. M. *et al.* A comprehensive survey of clonal diversity measures in Barrett's esophagus as biomarkers of progression to esophageal adenocarcinoma. *Cancer Prev. Res.* **3,** 1388–1397 (2010).
66. Dick, J. E. Stem cell concepts renew cancer research. *Blood* **112,** 4793–4807 (2008).
67. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414,** 105–111 (2001).
68. Greaves, M. Cancer stem cells renew their impact. *Nature Med.* **17,** 1046–1048 (2011).
69. Rosen, J. M. & Jordan, C. T. The increasing complexity of the cancer stem cell paradigm. *Science* **324,** 1670–1673 (2009).
70. Greaves, M. Cancer stem cells: back to Darwin? *Semin. Cancer Biol.* **20,** 65–70 (2010).
71. Gupta, P. B. *et al.* Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138,** 645–659 (2009).
72. Jamieson, C. H. *et al.* Granulocyte–macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. *N. Engl. J. Med.* **351,** 657–667 (2004).
73. Akala, O. O. *et al.* Long-term haematopoietic reconstitution by $Trp53^{-/-}p16^{Ink4a-/-}p19^{Arf-/-}$ multipotent progenitors. *Nature* **453,** 228–232 (2008).
74. Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442,** 818–822 (2006).
75. Olivier, M. & Taniere, P. Somatic mutations in cancer prognosis and prediction: lessons from *TP53* and *EGFR* genes. *Curr. Opin. Oncol.* **23,** 88–92 (2011).
76. Mizuno, H., Spike, B. T., Wahl, G. M. & Levine, A. J. Inactivation of *p53* in breast cancers correlates with stem cell transcriptional signatures. *Proc. Natl Acad. Sci. USA* **107,** 22745–22750 (2010).
77. Cicalese, A. *et al.* The tumor suppressor *p53* regulates polarity of self-renewing divisions in mammary stem cells. *Cell* **138,** 1083–1095 (2009).
78. Quintana, E. *et al.* Efficient tumour formation by single human melanoma cells. *Nature* **456,** 593–598 (2008).
**New xenograft methods revealed that cancer stem cells are common cell types in melanoma.**
79. Pece, S. *et al.* Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* **140,** 62–73 (2010).
80. Notta, F. *et al.* Evolution of human *BCR-ABL1* lymphoblastic leukaemia-initiating cells. *Nature* **469,** 362–367 (2011).
81. Clappier, E. *et al.* Clonal selection in xenografted human T cell acute

lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J. Exp. Med.* **208,** 653–661 (2011).
82. Frank, N. Y., Schatton, T. & Frank, M. H. The therapeutic promise of the cancer stem cell concept. *J. Clin. Invest.* **120,** 41–50 (2010).
83. Ishikawa, F. *et al.* Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nature Biotechnol.* **25,** 1315–1321 (2007).
84. Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* **1805,** 105–117 (2010).
85. Piccirillo, S. G. M. *et al.* Distinct pools of cancer stem-like cells coexist within human glioblastomas and display different tumorigenicity and independent genomic evolution. *Oncogene* **28,** 1807–1811 (2009).
86. Solit, D. & Sawyers, C. L. How melanomas bypass new therapy. *Nature* **468,** 902–903 (2010).
87. Goff, D. & Jamieson, C. Cycling toward elimination of leukemic stem cells. *Cell Stem Cell* **6,** 296–297 (2010).
88. Sharma, S. V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141,** 69–80 (2010).
89. Chin, L., Andersen, J. N. & Futreal, P. A. Cancer genomics: from discovery science to personalized medicine. *Nature Med.* **17,** 297–303 (2011).
90. Sawyers, C. L. Shifting paradigms: the seeds of oncogene addiction. *Nature Med.* **15,** 1158–1161 (2009).
91. Graham, S. M. *et al.* Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 *in vitro*. *Blood* **99,** 319–325 (2002).
92. Turke, A. B. *et al.* Preexistence and clonal selection of *MET* amplification in *EGFR* mutant NSCLC. *Cancer Cell* **17,** 77–88 (2010).
93. Ashworth, A., Lord, C. J. & Reis-Filho, J. S. Genetic interactions in cancer progression and treatment. *Cell* **145,** 30–38 (2011).
94. Zhang, B. *et al.* Effective targeting of quiescent chronic myelogenous leukemia stem cells by histone deacetylase inhibitors in combination with imatinib mesylate. *Cancer Cell* **17,** 427–442 (2010).
95. Duy, C. *et al.* *BCL6* enables Ph$^+$ acute lymphoblastic leukaemia cells to survive *BCR–ABL1* kinase inhibition. *Nature* **473,** 384–388 (2011).
96. Pienta, K. J., McGregor, N., Axelrod, R. & Axelrod, D. E. Ecological therapy for cancer: defining tumors using an ecosystem paradigm suggests new opportunities for novel cancer treatments. *Trans. Oncol.* **1,** 158–164 (2008).
97. Calabrese, C. *et al.* A perivascular niche for brain tumor stem cells. *Cancer Cell* **11,** 69–82 (2007).
98. Bissell, M. J. & Hines, W. C. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nature Med.* **17,** 320–329 (2011).
99. Wargo, A. R., Huijben, S., de Roode, J. C., Shepherd, J. & Read, A. F. Competitive release and facilitation of drug-resistant parasites after therapeutic chemotherapy in a rodent malaria model. *Proc. Natl Acad. Sci. USA* **104,** 19914–19919 (2007).
100. Gatenby, R. A., Silva, A. S., Gillies, R. J. & Frieden, B. R. Adaptive therapy. *Cancer Res.* **69,** 4894–4903 (2009).
**Dosing to maintain tumour size prolonged survival far longer than high-dose therapy in a mouse xenograft model.**

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** The research of M.G. is supported by Leukaemia & Lymphoma Research UK and The Kay Kendall Leukaemia Fund. The research of C.M. is supported by Research Scholar Grant #117209-RSG-09-163-01-CNE from the American Cancer Society and NIH grants P01 CA91955, U54 CA143803, R01 CA149566 and R01 CA140657. The authors thank C.Cooper and J.Clark for the use of the image in Fig. 4.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to M.G. (mel.greaves@icr.ac.uk).

doi:10.1038/nature10763

# The contribution of bone to whole-organism physiology

Gérard Karsenty[1] & Mathieu Ferron[1]

**The mouse genetic revolution has shown repeatedly that most organs have more functions than expected. This has led to the realization that, in addition to a molecular and cellular approach, there is a need for a whole-organism study of physiology. The skeleton is an example of how a whole-organism approach to physiology can broaden the functions of a given organ, reveal connections of this organ with others such as the brain, pancreas and gut, and shed new light on the pathogenesis of degenerative diseases affecting multiple organs.**

There are two definitions of physiology. In its most modern incarnation, physiology focuses on signalling and gene expression events occurring within cells as they relate to a given function and is often referred to as molecular and cellular physiology. An older physiology deals with events occurring and molecules acting outside the cell. It can be called whole-organism physiology and aims at identifying interactions taking place between organs through what Claude Bernard called the "*milieu intérieur*"[1]. Its study relies on three basic principles. First, the premise of whole-organism physiology is that no function is determined by one organ alone; this hypothesis has been verified multiple times through the use of model organisms[2–4]. Second, homeostasis, a principle positing that different organs exert opposite influences on the same function to regulate it tightly, applies to most physiological functions[5]. Closely related to homeostasis is the cardinal rule of endocrinology, that is feedback regulation: a regulated organ talks back to a regulating one to limit its influence[6]. Third is the principle in which whole-organism physiology resembles any other aspect of biology: regulatory molecules appear during evolution with the functions they regulate, not millions of years afterwards.

In vertebrates the modern study of whole-organism physiology uses two main tools. The experimental one, which favoured the renaissance of this physiology, is mouse genetics. This has allowed us to decipher, one gene at a time, how organs influence each other. For the focus of this Review, namely skeleton physiology, mouse genetics has proved to be an extremely reliable tool, possibly because bone is one of the last tissues to appear during evolution, and consequently most genes involved in skeletal biology have conserved functions from mice to humans. The second tool is the observation of internal medicine: mirror images of many physiological processes may be found in disease symptoms or drug side effects. Both tools are equally useful in extending the range of functions of the skeleton and our understanding of their molecular bases.

In this Review, we show how a whole-organism approach to physiology has modified our view of the skeleton. To that end, we describe work pointing towards a coordinated regulation of bone mass, energy metabolism and fertility; a central control of bone mass; and the roles of leptin, serotonin, insulin and osteocalcin in these pathways.

## The skeleton and whole-organism physiology

A strategy that could be used to determine whether the skeleton is influenced by organs that are not classically associated with it and whether it influences other organs is to confront features of bone that are unique to clinical and experimental observations. This approach may indicate physiological functions that could affect or be influenced by bone. The reality of such interactions could then be tested genetically.

A striking feature of bone is that it is the only tissue that contains a cell type, the osteoclast, whose unique function is to resorb (destroy) the host tissue[7]. This does not occur at random but rather in the context of a true homeostatic function, which is called bone modelling during childhood and remodelling during adulthood. This function, hereafter referred to as bone (re)modelling, is characterized by alternating phases of destruction by osteoclasts and bone formation by osteoblasts[8]. Bone modelling allows longitudinal growth, without which most vertebrates could not ambulate and, therefore, could not live. It is, by definition, a survival function. Bone (re)modelling occurs daily in multiple locations in an organ covering a very large surface area. Both the cellular events it entails and the surface area of the organ in which they occur suggest there is a high energetic cost. Clinical observations add credibility to this view of bone (re)modelling as an energy-demanding process. Specifically, the absence of food — that is, energy — intake, as in patients with anorexia nervosa, causes a near-total arrest of growth in children and low bone mass in adults[9,10]. Moreover, in a manner unrelated to food intake, gonadal failure leads to low bone mass in both sexes, thereby suggesting a link between bone mass accrual and fertility[11,12]. Although compelling, such clinical evidence remained correlative. It became an incentive for laboratory investigation because of an experimental observation that was striking because it was unexpected. Osteocalcin was, at the time the encoding gene was inactivated in the mouse, the only osteoblast-specific secreted protein[13]. This by itself justified the study of its function *in vivo* with the hope of learning more about bone biology. Surprisingly, osteocalcin deletion resulted in mice that were abnormally obese and that bred poorly[14,15]. These phenotypes suggested that, in ways that still need to be explained, bone was affecting fat accumulation and possibly other aspects of energy metabolism and reproduction.

Taken together, this view of bone (re)modelling, clinical observations and the phenotypes of mice deficient in the osteocalcin gene suggested that there might have been a coordinated regulation of bone mass or growth, energy metabolism and reproduction. Because this hypothesis was triggered by the energy cost of bone (re)modelling, one would expect that the hormones orchestrating it would appear with the skeleton during evolution. Several laboratories have tested this hypothesis in the past 10 years.

---

[1] Department of Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA.

## A whole-organism physiology view of leptin

For a bone biologist in the late 1990s, it seemed easier to explore the aforementioned assumption by studying a hormone for which the biology has been characterized and the receptor identified. This hormone was leptin, an adipocyte-specific molecule identified for its ability to limit appetite and to favour energy expenditure and reproduction, two tenets of the overarching hypothesis[16–18]. Remarkably, and this is important in view of its function described here, leptin does not appear during evolution with any aspect of energy metabolism or reproduction, but is instead associated with bone (re)modelling. This would be mere coincidence if *ob*/*ob* (obese) or *db*/*db* (diabetic) mice, which lack leptin or its receptor, respectively, did not demonstrate a high bone mass because of a massive increase in bone formation[19]. This is a biological tour de force because these mice have no functional gonads, a situation that would otherwise increase bone resorption and decrease bone mass. Subsequently, leptin regulation of bone mass was verified in sheep and humans[20–22]. The high bone mass (despite hypogonadism) observed in the absence of leptin signalling and the fact that leptin is a vertebrate invention suggested that bone was a major target of this hormone; a model of partial gain-of-function of leptin signalling allowed this hypothesis to be tested. To mediate its functions, leptin binds to a receptor (Lepr) linked to the tyrosine kinase Jak2. Leptin binding activates Jak2, resulting in the phosphorylation of several residues on Lepr. One of these, Tyr 985, binds Socs3, an event that attenuates signalling through Lepr[23] (Fig. 1). Accordingly, mutation of this residue in Lepr (Y985L) in *l*/*l* mice, which are homozygous for this substitution, results in a partial gain of function of leptin signalling[24]. Because the increase in signalling is only partial, phenotypes displayed by *l*/*l* mice reveal one or more functions of leptin that require its lowest threshold of signalling[24]. Hence, *l*/*l* mice breed normally and have normal appetite when fed on normal chow, but are osteoporotic; these observations suggest that the threshold of leptin signalling necessary to affect bone mass is lower than that needed to affect appetite and reproduction[25].

Several genetic pieces of evidence indicate that leptin acts centrally to inhibit the accrual of bone mass. The most convincing argument, although not the only one, is that a neuron-specific deletion of *Lepr* recapitulates the bone phenotype of *ob*/*ob* mice, whereas an osteoblast-specific one does not[25]. This is consistent with what has been shown for other functions of leptin that also occur through a central relay. In broader terms this revealed for the first time a central control of bone mass; its existence has now been verified by other laboratories studying how neuropeptide Y or neuromedin U regulate bone mass[26,27]. Over the years it has been proposed that, in contrast to most of its functions, leptin could also regulate bone mass through local means. This hypothesis is based on studies that injected large amounts (micrograms per day) of leptin into wild-type mice[28] or infused leptin into the third ventricle of *ob*/*ob* mice[29]. That leptin did not decrease the percentage of fat in the latter study, in contrast to expectations, raises questions about the efficacy of these infusions.

Let us concentrate here on the central mode of action of leptin. We needed to know where it signals in the brain to fulfil this function and the regulation of energy metabolism (we viewed these two functions as co-regulated). Initially, it seemed that the hypothalamus was where all the action was taking place. The leptin receptor is highly expressed in ventromedial (VMH) and arcuate neurons of the hypothalamus, and chemical lesioning of these neurons had demonstrated their involvement in the regulation of appetite and bone mass; moreover, leptin infusions in the third ventricle of *ob*/*ob* mice decreased bone mass and appetite only if these hypothalamic neurons were intact[30]. This simple view was shattered by landmark studies showing that the selective inactivation of *Lepr* in VMH or arcuate neurons did not affect appetite or the accrual of bone mass in mice fed on a normal diet, the diet on which *ob*/*ob* mice display hyperphagia[31–33]. One interpretation of these seemingly contradictory results is that they are instead complementary in suggesting that leptin requires the integrity of hypothalamic neurons to regulate appetite and bone mass, but it need not bind to them. In other words, leptin might signal elsewhere in the brain to regulate the synthesis of neurotransmitters that will then act in the hypothalamus. Admittedly, this is a shift in the view of leptin signalling in the brain, but a novel approach was needed because cell-specific deletion of *Lepr* in hypothalamic neurons did not reproduce phenotypes seen in *ob*/*ob* mice fed on normal chow. The fact that patients chronically treated with serotonin reuptake inhibitors can develop low bone mass suggested that serotonin and leptin signalling intersect in the brain[34–36].

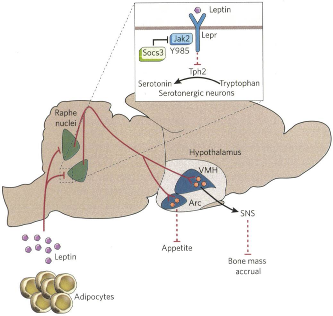


**Figure 1 | Leptin co-regulates appetite and bone mass.** Leptin signals through its receptor (Lepr), expressed in serotonin-producing neurons of the raphe nuclei in the brainstem, decreasing the expression of tryptophan hydrxylase 2 (*Tph2*), the gene encoding the initial enzyme for serotonin biosynthesis. Signalling by Lepr requires Jak2 and is negatively regulated by Socs3. Serotonin-producing neurons of the raphe nuclei project to the ventromedial (VMH) and the arcuate (Arc) hypothalamic nuclei. Serotonin signalling in VMH neurons decreases the activity of the sympathetic nervous system (SNS), an inhibitor of bone mass accrual. In Arc neurons, serotonin signalling has anti-anorexigenic effects.

In some instances cell-specific gene inactivation can be achieved without using sophisticated techniques. These are favourable circumstances because cell-specific gene deletion techniques are not without risks. This situation presented itself for leptin and for brain-derived serotonin, which is made only in neurons of the raphe nuclei and does not cross the blood–brain barrier[37]. It could therefore be eliminated from the brain by conventional inactivation of tryptophan hydroxylase 2 (Tph2), the rate-limiting enzyme of serotonin synthesis[38]. The fact that *Tph2*$^{-/-}$ mice were distinctly osteoporotic and anorectic established that brain-derived serotonin favours bone mass accrual and appetite and, because it does not cross the blood–brain barrier, identified serotonin as the first neurotransmitter regulating bone mass accrual[31]. Axonal connections between serotonin-producing neurons and hypothalamic neurons can exist and are probably functionally important, because ablations of serotonin receptors in specific neurons of the hypothalamus result in osteoporosis or anorexia in the same way as *Tph2* deletion does[31,39,40]. This work, suggesting a direct brainstem–hypothalamus axis, does not contradict or exclude a more recent study proposing that serotonin signalling in the hypothalamus occurs through an interneuron[41]. This interneuron will need to be identified for the proper evaluation of the respective importance of these two modes of action.

Leptin inhibits *Tph2* expression in, and serotonin release from, brainstem neurons; consequently, serotonin content in *ob/ob* hypothalami is abnormally high. Removing one allele of *Tph2* from *ob/ob* mice sufficed to normalize their brain serotonin content, with remarkable consequences[31]. Indeed, *ob/ob*;$Tph2^{+/-}$ mice, despite having no leptin, had normal appetite, energy expenditure, body weight and bone mass. Subsequently, the molecular bases of this leptin–serotonin–hypothalamus axis were deciphered. They include calmodulin signalling in VMH neurons activating the transcription factor cyclic-AMP-response-element-binding protein (CREB), which induces the expression of genes involved in catecholamine synthesis[39]. This is important because the mediator acting as a bridge between leptin signalling in the brain and in bone cells is the sympathetic nervous system acting on osteoblasts through the $\beta_2$ adrenergic receptor (Adrb2)[30,42] (Fig. 1) and because β-blockers can limit the risk of osteoporotic fractures[43].

The fact that leptin inhibits both bone mass accrual and appetite is consistent with the hypothesis that bone acquisition, an energy-expensive process, must be linked to energy (food) intake, otherwise the risk of organ failure anywhere else in the body, at the time of a growth spurt for instance, could be high. This broader view of leptin proposes a simple explanation of why this hormone appeared during evolution with bone and when food was scarce.

## Osteocalcin and the endocrine nature of bone

Implicit in the hypothesis that there is coordinated regulation of bone mass, energy metabolism and fertility is the notion that bone is not only a recipient of hormonal inputs but also an endocrine organ affecting the other two functions. It was already known that, through fibroblast growth factor 23 (FGF 23), bone acts as an endocrine organ, but the function regulated by FGF23, namely mineral metabolism, is intimately linked to bone health[44,45]. The question asked at this point was different: does bone regulate physiological functions that, a priori, have nothing to do with bone health?

Although this hypothesis was initially formulated on inspection of mice deficient in the osteocalcin gene, it was the study of another gene and of another mouse model that shaped the concept into the form of a genetic pathway. Embryonic stem-cell phosphatase (ESP) is an obscure protein tyrosine phosphatase without any known substrate that is expressed in only three cell types: the embryonic stem cell, the Sertoli cell of the testes and the osteoblast[15]. This restricted pattern of expression justified the study of its function *in vivo*, and for that purpose *Esp* was deleted either in all cells or only in osteoblasts. The observation that both mouse models developed the same phenotypes implied that the biological functions they reveal take place in osteoblasts[15].

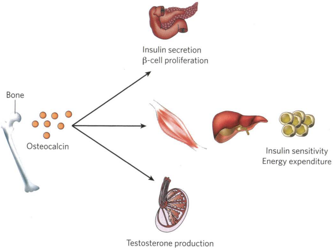


**Figure 2 | Osteocalcin, a bone-derived multifunctional hormone.** Undercarboxylated osteocalcin stimulates insulin secretion and β-cell proliferation in the pancreas, energy expenditure by muscle, and insulin sensitivity in adipose tissue, muscle and liver. In addition, it promotes male fertility by stimulating testosterone synthesis in Leydig cells of the testis through the activation of its receptor, GPRC6A, in these cells.

The absence of *Esp* in osteoblasts prompted perinatal death of mice, resulting from severe hypoglycaemia. This happened because the functions of this intracellular tyrosine phosphatase are, through its expression in osteoblasts, to inhibit the expression and secretion of insulin by pancreatic β-cells, to limit insulin sensitivity in liver, muscle and white adipose tissue, and to decrease energy expenditure[15]. All of the phenotypes of *Esp*-deficient mice mirrored those observed in mice deficient in the osteocalcin gene, and genetically *Esp* inhibits osteocalcin's functions[15]. The fact that wild-type, but not osteocalcin-gene-deficient, osteoblasts induce insulin secretion from isolated islets established that osteocalcin is a hormone promoting β-cell proliferation and insulin expression and secretion. It also increases insulin sensitivity and energy expenditure[15] (Fig. 2).

Before considering other aspects of osteocalcin biology one needs to address the relevance of these findings to human physiology. To the best of our knowledge no molecule identified as a hormone in the mouse has lost this function in humans. Although no mutation in osteocalcin or its receptor has yet been reported, this assumption seems to extend to osteocalcin, which has become an increasingly accepted biomarker of insulin resistance in human studies[46,47].

Because most hormones have several functions, the next question was whether this was the case for osteocalcin. This was even more relevant because osteocalcin-gene-deficient mice breed poorly. A study of this phenotype revealed that osteocalcin promotes testosterone synthesis by Leydig cells of the testis and fertility in male mice[14]. For that purpose osteocalcin upregulates, in testes but not in ovaries, the synthesis of enzymes necessary for testosterone biosynthesis, without affecting the expression of the gene encoding Cyp19 (the enzyme converting testosterone to oestradiol). This male-specific function of osteocalcin implied that its putative receptor (OSTR) may be expressed only in testes. This sexual dichotomy allowed the identification of GPRC6A, a G-protein-coupled receptor, as an osteocalcin receptor, that is expressed in mice and humans in Leydig cells of the testis, but not in follicular cells of the ovary[14] (Fig. 2). A pioneering study had previously proposed that GPR6CA could be an osteocalcin receptor, although no binding experiments were done[48].

## Osteocalcin puts bone in the thick of things

If the notion that bone affects energy metabolism and reproduction was implicit in the overall hypothesis, what emerged next — that bone is an endocrine hub on which several hormones converge to recruit osteocalcin as their ultimate relay — was not expected. To appreciate this notion, which was an outcome of the study of the relationship between ESP and osteocalcin, one needs to understand how osteocalcin is modified post-translationally.

Osteocalcin is carboxylated on three glutamic acid residues, a modification that confers on proteins a high affinity for minerals such as the hydroxyapatite crystal present in the mineralized bone matrix[13]. Nevertheless, a small but measurable proportion of undercarboxylated osteocalcin is found in the serum[49], indicating that either some osteocalcin is secreted in an incompletely carboxylated form by osteoblasts or that osteocalcin becomes decarboxylated outside the cell. This was an important question to address because cell-based and *in vivo* investigations had shown that it is the undercarboxylated form of osteocalcin (the one in which Glu 13 is not carboxylated) that is active on β-cells and Leydig cells[14,50].

The search for substrates of ESP identified the insulin receptor as such a molecule[51]. The *Esp*-null mouse, therefore, is a gain-of-function model of insulin signalling in osteoblasts. Subsequently, two groups working independently showed that mice lacking the insulin receptor in osteoblasts ($InsR_{osb}^{-/-}$ mice) were glucose intolerant and insulin insensitive when fed on normal chow; that is, they were a phenocopy of the osteocalcin-gene-deficient mice[51,52]. Because mice lacking the insulin receptor in skeletal muscle or white adipose tissue do not display glucose intolerance when fed on a normal diet[53,54], insulin must act in additional tissues to achieve glucose homeostasis. The fact that bone is such a tissue legitimizes the notion that this tissue is necessary for glucose homeostasis. In addition, $InsR_{osb}^{-/-}$ mice had significantly less biologically active (undercarboxylated) osteocalcin in their sera, revealing that insulin signalling in osteoblasts is a determinant of osteocalcin bioactivity[53,54]. In a manner that is both elegant and economical, insulin uses the interplay between osteoblasts and osteoclasts for that purpose. Specifically, insulin inhibits the expression in osteoblasts of the gene encoding osteoprotegerin (*Opg*)[51], which hampers osteoclast differentiation. In other words, insulin signalling in osteoblasts favours bone resorption, a process that occurs at pH 4.5 (ref. 55). Acidic pH is the only mechanism known to achieve decarboxylation of proteins[56], therefore, bone resorption decarboxylates and activates osteocalcin[51]. Thus, in a feedforward loop, insulin signalling in osteoblasts promotes its own secretion by activating osteocalcin (Fig. 3), and mice and humans in which bone resorption is genetically impaired show a decrease in the undercarboxylated form of osteocalcin, resulting in glucose intolerance[51]. The functional equivalent of ESP in human osteoblasts is protein tyrosine phosphatase 1B (PTP-1B), a tyrosine phosphatase previously known for its ability to inactivate the insulin receptor in other cell types[51,57–59].

If insulin enhances osteocalcin activity, which in turn favours insulin secretion, how do vertebrates prevent hypoglycaemia? The answer is leptin[60]. Young *ob*/*ob* mice, before they become obese, have no molecular or metabolic evidence of insulin resistance, yet they are markedly hyperinsulinaemic and hypoglycaemic, indicating that a function of leptin is to inhibit insulin secretion[3]. As is true of its other functions, leptin by and large does not act locally on islets to affect insulin secretion, but through a neuronal relay[60]. Because the sympathetic nervous system acting through Adrb2 expressed in osteoblasts mediates the inhibition of bone mass accrual by leptin, it was tempting to ask whether the same pathway might inhibit insulin secretion by modulating osteocalcin expression or activity. Molecular evidence showed that the sympathetic tone mediates the upregulation of *Esp* expression by leptin and that *ob*/*ob* mice have more active osteocalcin in their circulation[60]. Three pieces of genetic evidence demonstrate that leptin regulation of osteocalcin activity occurs through sympathetic signalling in osteoblasts. First, mice lacking *Adrb2* in osteoblasts are hypoglycaemic and hyperinsulinaemic. Second, mice lacking one copy of the leptin gene and one copy of *Adrb2* in osteoblasts are also hyperinsulinaemic and hypoglycaemic. Third, *ob*/*ob* mice that lack the osteocalcin gene have normal expression of the insulin genes and are normo-insulinaemic for significantly longer than *ob*/*ob* mice[60].

Altogether, the metabolic functions of osteocalcin and their regulation by leptin and insulin reveal how intertwined energy metabolism and bone (re)modelling are. They also underscore the importance of the interplay between osteoblasts and osteoclasts for glucose metabolism and indicate that the regulation of osteocalcin activity occurs at both the transcriptional and post-translational levels.

## Energy metabolism, bone mass and bone diseases

Energy metabolism is a multistep process initiated by food absorption along the gastrointestinal tract. Thus, in considering whether there is coordinated regulation of bone mass and energy metabolism one needs to determine whether any part or function of the gastrointestinal tract influences the accrual of bone mass. This is important in view of the fact that the gastrointestinal tract is an endocrine organ secreting several hormones whose functions are not all known. Clinically there is reason to ask whether the gastrointestinal tract affects bone mass: osteoporosis is often seen in patients with inflammatory bowel diseases[61,62].

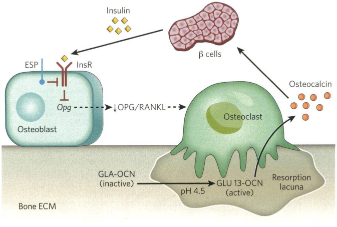


**Figure 3 | A feedforward loop links insulin, bone resorption and osteocalcin activity.** Insulin signalling in osteoblasts decreases the expression of *Opg*. The decrease in the ratio of osteoprotegerin (OPG) to receptor activator of nuclear factor-κB ligand (RANKL) increases bone resorption by osteoclasts. The acidic pH (4.5) in resorption lacunae decarboxylates (that is, activates) osteocalcin (GLA-OCN) stored in the bone extracellular matrix. Undercarboxylated active osteocalcin (GLU13-OCN) then stimulates insulin secretion by the β-cells of the pancreatic islets and promotes insulin sensitivity in peripheral organs. ECM, extracellular matrix; InsR, insulin receptor.

The study of mouse models of two human genetic diseases, the Coffin–Lowry syndrome (a learning disability syndrome with low bone mass) and the skeletal manifestations of neurofibromatosis 1 (NF1), established the existence of a connection between food absorption and the accrual of bone mass. In these disorders, bone mass abnormalities are due to a decrease or an increase, respectively, in the activity of activating transcription factor 4 (ATF4), a transcription factor that favours bone formation, in part by promoting amino acid import into osteoblasts[63,64]. One could, therefore, put this function of ATF4 to work to correct the skeletal manifestations of Coffin–Lowry syndrome or NF1 in the mouse by modulating protein content in their diets[63] (Fig. 4). Another study looked at the influence of the gastrointestinal tract on mineral metabolism. Given that bone is a mineralized tissue and that the mineral crystal in the bone is made of calcium and phosphate, it is not surprising that significant variations in calcium or phosphate metabolism affect bone mass. One reason for that is that a high extracellular concentration of calcium inhibits secretion by parathyroid glands of parathyroid hormone (PTH), whose physiological role is to increase bone resorption[65]. A study relying on mouse genetics and human pathology showed that the high acidity (low pH) present in the stomach, the entry point of the gastrointestinal tract, is required for proper calcium absorption and, therefore, for normal PTH secretion and bone resorption[66] (Fig. 4). This established a simple connection between the gastrointestinal tract and bone mass that could be exploited for therapeutic purposes. These examples of an influence of the gastrointestinal tract on bone mass raised the possibility that hormones made by the gastrointestinal tract regulate bone mass. A peculiar combination of scientific frustration and luck showed that this was not only the case, but also that this regulation is potentially important for the treatment of the most frequent bone degenerative disease, osteoporosis.

The cell-surface molecule Lrp5 (low-density-lipoprotein receptor related protein 5), despite having no identifiable ligand and no signalling pathway, is remarkably important in bone biology for medical reasons. First, *LRP5* is mutated in two human diseases. In osteoporosis pseudoglioma (OPPG), a disease characterized by the appearance of osteoporosis several years after birth, it is inactivated[67]; in high bone mass syndrome (HBM) there is a missense mutation in *LRP5* that is thought to be a gain-of-function mutation[68,69]. Patients

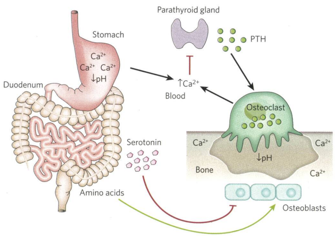


**Figure 4 | Interactions between the gastrointestinal tract and bone mass.** Stomach acidity (low pH) is required for the proper absorption of calcium ($Ca^{2+}$) and is, therefore, essential to maintain normal levels of serum calcium. Serum calcium, in turn, negatively regulates secretion from the parathyroid gland of PTH, a hormone that stimulates osteoclast differentiation and bone resorption. Bone resorption by osteoclasts also occurs at low pH and contributes to the maintenance of serum calcium. Peripheral serotonin is produced by the duodenum and inhibits bone formation by osteoblasts, whereas dietary intake of amino acids (proteins) favours collagen synthesis by osteoblasts.

harbouring the latter mutation have no detectable symptoms before adulthood, again arguing against a developmental defect. In fact, the main symptom is that post-menopausal female patients with HBM do not develop osteoporosis[68,69]. Another reason to draw much attention to Lrp5 is that it affects bone mass only by acting on bone formation[70]. In other words, the Lrp5-dependent signalling pathway holds the key to an anabolic treatment for osteoporosis, the ultimate objective for the most frequent bone degenerative disease.

Lrps are known to be able to bind multiple ligands; nevertheless, because of its homology to Arrow[71] (a co-receptor in *Drosophila* for Wingless, a homologue of Wnt) it is as a potential co-receptor for Wnt proteins that Lrp5 has been studied for the longest time. This was a reasonable assumption to test, and cell culture experiments showed repeatedly that Lrp5, like other Lrps, can function as a Wnt co-receptor *in vitro* and in controlling eye vascularization[71–75]. However, when it came to its role in bone, there were discrepancies between expectations and results obtained through the study of animal models. A few examples follow. *Lrp5*-null osteoblasts proliferate poorly *in vivo,* but normally in cell culture[76]. Because Lrp5 is a receptor, this observation is inconsistent with the notion that its signalling is affected in osteoblasts. Furthermore, microarray analyses showed that genes regulated by Lrp5 (cell-cycle regulators) and by the canonical Wnt signalling (*Opg*) in bone specimens were different[77,78]. Accordingly, in mice, inactivation of the canonical Wnt signalling pathway in osteoblasts or osteocytes does not affect bone formation but bone resorption instead[78,79]. In humans, patients with HBM who have been followed up for more than 20 years have not yet developed bone tumours (M. Kassem, personal communication), even though many Wnt family members are aggressive oncoproteins[80].

In an unforeseen turn of events, attempts to understand these discrepancies have brought serotonin back into the picture. While studying the role of brain-derived serotonin through the inactivation of *Tph2*, the function of Tph1, the enzyme responsible for peripheral serotonin synthesis, was investigated as a negative control. *Tph1*$^{-/-}$ mice had already been generated when it was realized that the most overexpressed gene in the absence of *Lrp5* was in fact *Tph1* (ref. 76). From being an internal control in another project *Tph1* became, overnight, the topic of an independent one. Two groups have generated mouse models of *Lrp5* deletion in a cell-specific manner and have obtained different results[76,81]. The reason for this discrepancy is unknown. Because we ourselves generated one of these models, we believe it would be inappropriate to use the tribune provided by this review to make our case. It is also unnecessary, because, as important as mouse genetics is, it remains a surrogate for human genetics, and so far all patients with OPPG or HBM for whom measurements of circulating serotonin have been published have shown high and low circulating levels of serotonin, respectively[76,82–84]. Similarly, serotonin-producing tumours cause osteoporosis[85], and patients with osteoporosis have high circulating serotonin levels[86] (Fig. 4). As a result, the regulation of bone mass by serotonin has now acquired a (medical) life of its own and deserves to be studied because it holds great promise for the anabolic treatment of osteoporosis. This is even more true given that two proof-of-principle studies have shown that an inhibitor of serotonin synthesis in the gut, shown to be safe in humans in a phase I clinical trial[87,88], cured bone disease in *Lrp5*-null mice and gonadectomy-induced osteoporosis in rodents. In both cases this was achieved through a purely anabolic mode of action, just as Lrp5 affects bone mass[89,90].

## Perspective

The most appropriate way to look forward is to formulate the questions that will identify the next frontiers in this aspect of physiology. For instance, are all of the functions of osteocalcin known? What are the genes downstream of osteocalcin in all of its target cells? Is osteocalcin the only bone-derived hormone affecting energy metabolism? Is there a counterpart of osteocalcin in females that affects their reproduction? We will also venture a general and humbling statement: if we have learned so much in so brief a time about the physiology of one organ, it is likely that there are many aspects of whole-organism physiology to be discovered through the use of model organisms. ■

1. Bernard, C. *Introduction à l'Etude de la Médecine Expérimentale* (Flammarion, 1865).
2. Gurney, C. W. Erythropoietin, erythropoiesis, and the kidney. *J. Am. Med. Assoc.* **173,** 1828–1829 (1960).
3. Friedman, J. M. & Halaas, J. L. Leptin and the regulation of body weight in mammals. *Nature* **395,** 763–770 (1998).
4. Chen, S. K. *et al.* Hematopoietic origin of pathological grooming in *Hoxb8* mutant mice. *Cell* **141,** 775–785 (2010).
5. Cannon, W. B. *The Wisdom of the Body* (Norton, 1932).
6. Monod, J. & Jacob, F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26,** 389–401 (1961).
7. Teitelbaum, S. L. Bone resorption by osteoclasts. *Science* **289,** 1504–1508 (2000).
8. Ducy, P., Schinke, T. & Karsenty, G. The osteoblast: a sophisticated fibroblast under central surveillance. *Science* **289,** 1501–1504 (2000).
9. Legroux-Gerot, I., Vignau, J., Collier, F. & Cortet, B. Bone loss associated with anorexia nervosa. *Joint Bone Spine* **72,** 489–495 (2005).
10. Misra, M. & Klibanski, A. The neuroendocrine basis of anorexia nervosa and its impact on bone metabolism. *Neuroendocrinology* **93,** 65–73 (2011).
11. Riggs, B. L., Khosla, S. & Melton, L. J. A unitary model for involutional osteoporosis: estrogen deficiency causes both type I and type II osteoporosis in postmenopausal women and contributes to bone loss in aging men. *J. Bone Miner. Res.* **13,** 763–773 (1998).
12. Riggs, B. L. & Melton, L. J. Involutional osteoporosis. *N. Engl. J. Med.* **314,** 1676–1686 (1986).
13. Hauschka, P. V., Lian, J. B., Cole, D. E. & Gundberg, C. M. Osteocalcin and matrix Gla protein: vitamin K-dependent proteins in bone. *Physiol. Rev.* **69,** 990–1047 (1989).
14. Oury, F. *et al.* Endocrine regulation of male fertility by the skeleton. *Cell* **144,** 796–809 (2011).
15. Lee, N. K. *et al.* Endocrine regulation of energy metabolism by the skeleton. *Cell* **130,** 456–469 (2007).
**This paper describes for the first time the endocrine function of bone.**
16. Zhang, Y. *et al.* Positional cloning of the mouse obese gene and its human homologue. *Nature* **372,** 425–432 (1994).
**This report describes the identification of the leptin gene.**
17. Halaas, J. L. *et al.* Weight-reducing effects of the plasma protein encoded by the obese gene. *Science* **269,** 543–546 (1995).
18. Chehab, F. F., Lim, M. E. & Lu, R. Correction of the sterility defect in homozygous obese female mice by treatment with the human recombinant leptin. *Nature Genet.* **12,** 318–320 (1996).
19. Ducy, P. *et al.* Leptin inhibits bone formation through a hypothalamic relay:

a central control of bone mass. *Cell* **100,** 197–207 (2000).

20. Pogoda, P. *et al.* Leptin inhibits bone formation not only in rodents, but also in sheep. *J. Bone Miner. Res.* **21,** 1591–1599 (2006).
21. Elefteriou, F. *et al.* Serum leptin level is a regulator of bone mass. *Proc. Natl Acad. Sci. USA* **101,** 3258–3263 (2004).
22. Gibson, W. T. *et al.* Congenital leptin deficiency due to homozygosity for the Δ133G mutation: report of another case and evaluation of response to four years of leptin therapy. *J. Clin. Endocrinol. Metab.* **89,** 4821–4826 (2004).
23. Bjorbaek, C., Elmquist, J. K., Frantz, J. D., Shoelson, S. E. & Flier, J. S. Identification of SOCS-3 as a potential mediator of central leptin resistance. *Mol. Cell* **1,** 619–625 (1998).
24. Bjornholm, M. *et al.* Mice lacking inhibitory leptin receptor signals are lean with normal endocrine function. *J. Clin. Invest.* **117,** 1354–1360 (2007).
25. Shi, Y. *et al.* Dissociation of the neuronal regulation of bone mass and energy metabolism by leptin *in vivo*. *Proc. Natl Acad. Sci. USA* **105,** 20529–20533 (2008).
26. Baldock, P. A. *et al.* Neuropeptide Y knockout mice reveal a central role of NPY in the coordination of bone mass to body weight. *PLoS ONE* **4,** e8415 (2009).
27. Sato, S. *et al.* Central control of bone remodeling by neuromedin U. *Nature Med.* **13,** 1234–1240 (2007).
28. Cornish, J. *et al.* Leptin directly regulates bone cell function *in vitro* and reduces bone fragility *in vivo*. *J. Endocrinol.* **175,** 405–415 (2002).
29. Bartell, S. M. *et al.* Central (ICV) leptin injection increases bone formation, bone mineral density, muscle mass, serum IGF-1, and the expression of osteogenic genes in leptin-deficient *ob/ob* mice. *J. Bone Miner. Res.* **26,** 1710–1720 (2011).
30. Takeda, S. *et al.* Leptin regulates bone formation via the sympathetic nervous system. *Cell* **111,** 305–317 (2002).
    **This paper reports that leptin regulation of bone mass requires the sympathetic nervous system.**
31. Yadav, V. K. *et al.* A serotonin-dependent mechanism explains the leptin regulation of bone mass, appetite, and energy expenditure. *Cell* **138,** 976–989 (2009).
    **This study identifies brain serotonin as a critical mediator of the central action of leptin.**
32. Balthasar, N. *et al.* Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron* **42,** 983–991 (2004).
33. Dhillon, H. *et al.* Leptin directly activates SF1 neurons in the VMH, and this action by leptin is required for normal body-weight homeostasis. *Neuron* **49,** 191–203 (2006).
34. Warden, S. J. *et al.* Psychotropic drugs have contrasting skeletal effects that are independent of their effects on physical activity levels. *Bone* **46,** 985–992 (2010).
35. Bliziotes, M. Update in serotonin and bone. *J. Clin. Endocrinol. Metab.* **95,** 4124–4132 (2010).
36. Kaye, W., Gendall, K. & Strober, M. Serotonin neuronal function and selective serotonin reuptake inhibitor treatment in anorexia and bulimia nervosa. *Biol. Psychiatry* **44,** 825–838 (1998).
37. Mann, J. J. *et al.* Relationship between central and peripheral serotonin indexes in depressed and suicidal psychiatric inpatients. *Arch. Gen. Psychiatry* **49,** 442–446 (1992).
38. Walther, D. J. *et al.* Synthesis of serotonin by a second tryptophan hydroxylase isoform. *Science* **299,** 76 (2003).
    **This report describes the cloning of the *Tph2* gene, which is responsible for brain serotonin synthesis.**
39. Oury, F. *et al.* CREB mediates brain serotonin regulation of bone mass through its expression in ventromedial hypothalamic neurons. *Genes Dev.* **24,** 2330–2342 (2010).
40. Yadav, V. K. *et al.* Leptin-dependent serotonin control of appetite: temporal specificity, transcriptional regulation, and therapeutic implications. *J. Exp. Med.* **208,** 41–52 (2011).
41. Lam, D. D. *et al.* Leptin does not directly affect CNS serotonin neurons to influence appetite. *Cell Metab.* **13,** 584–591 (2011).
42. Elefteriou, F. *et al.* Leptin regulation of bone resorption by the sympathetic nervous system and CART. *Nature* **434,** 514–520 (2005).
43. Bonnet, N. *et al.* Protective effect of beta blockers in postmenopausal women: influence on fractures, bone density, micro and macroarchitecture. *Bone* **40,** 1209–1216 (2007).
44. ADHR Consortium. Autosomal dominant hypophosphataemic rickets is associated with mutations in *FGF23*. *Nature Genet.* **26,** 345–348 (2000).
45. Hori, M., Shimizu, Y. & Fukumoto, S. Fibroblast growth factor 23 in phosphate homeostasis and bone metabolism. *Endocrinology* **152,** 4–10 (2011).
46. Motyl, K. J., McCabe, L. R. & Schwartz, A. V. Bone and glucose metabolism: a two-way street. *Arch. Biochem. Biophys.* **503,** 2–10 (2010).
47. Pollock, N. K. *et al.* Lower uncarboxylated osteocalcin concentrations in children with prediabetes is associated with beta-cell function. *J. Clin. Endocrinol. Metab.* **96,** E1092–E1099 (2011).
48. Pi, M. *et al.* Identification of a novel extracellular cation-sensing G-protein-coupled receptor. *J. Biol. Chem.* **280,** 40201–40209 (2005).
49. Merle, B. & Delmas, P. D. Normal carboxylation of circulating osteocalcin (bone Gla-protein) in Paget's disease of bone. *Bone Miner.* **11,** 237–245 (1990).
50. Ferron, M., Hinoi, E., Karsenty, G. & Ducy, P. Osteocalcin differentially regulates beta cell and adipocyte gene expression and affects the development of metabolic diseases in wild-type mice. *Proc. Natl Acad. Sci. USA* **105,** 5266–5270 (2008).
51. Ferron, M. *et al.* Insulin signaling in osteoblasts integrates bone remodeling and energy metabolism. *Cell* **142,** 296–308 (2010).
    **This study identifies a feed-forward loop linking insulin and osteocalcin.**
52. Fulzele, K. *et al.* Insulin receptor signaling in osteoblasts regulates postnatal bone acquisition and body composition. *Cell* **142,** 309–319 (2010).
    **In this report, insulin signalling in osteoblasts is identified as a key regulator of bone mass and of energy metabolism.**
53. Bluher, M. *et al.* Adipose tissue selective insulin receptor knockout protects against obesity and obesity-related glucose intolerance. *Dev. Cell* **3,** 25–38 (2002).
54. Bruning, J. C. *et al.* A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. *Mol. Cell* **2,** 559–569 (1998).
55. Silver, I. A., Murrills, R. J. & Etherington, D. J. Microelectrode studies on the acid microenvironment beneath adherent macrophages and osteoclasts. *Exp. Cell Res.* **175,** 266–276 (1988).
56. Poser, J. W. & Price, P. A. A method for decarboxylation of γ-carboxyglutamic acid in proteins. Properties of the decarboxylated γ-carboxyglutamic acid protein from calf bone. *J. Biol. Chem.* **254,** 431–436 (1979).
57. Elchebly, M. *et al.* Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* **283,** 1544–1548 (1999).
58. Delibegovic, M. *et al.* Liver-specific deletion of protein-tyrosine phosphatase 1B (PTP1B) improves metabolic syndrome and attenuates diet-induced endoplasmic reticulum stress. *Diabetes* **58,** 590–599 (2009).
59. Delibegovic, M. *et al.* Improved glucose homeostasis in mice with muscle-specific deletion of protein-tyrosine phosphatase 1B. *Mol. Cell. Biol.* **27,** 7727–7734 (2007).
60. Hinoi, E. *et al.* The sympathetic tone mediates leptin's inhibition of insulin secretion by modulating osteocalcin bioactivity. *J. Cell Biol.* **183,** 1235–1242 (2008).
61. Miheller, P., Lorinczy, K. & Lakatos, P. L. Clinical relevance of changes in bone metabolism in inflammatory bowel disease. *World J. Gastroenterol.* **16,** 5536–5542 (2010).
62. Rodriguez-Bores, L., Barahona-Garrido, J. & Yamamoto-Furusho, J. K. Basic and clinical aspects of osteoporosis in inflammatory bowel disease. *World J. Gastroenterol.* **13,** 6156–6165 (2007).
63. Elefteriou, F. *et al.* ATF4 mediation of *NF1* functions in osteoblast reveals a nutritional basis for congenital skeletal dysplasiae. *Cell Metab.* **4,** 441–451 (2006).
64. Yang, X. *et al.* ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin–Lowry syndrome. *Cell* **117,** 387–398 (2004).
65. Raisz, L. G. Stimulation of bone resorption by parathyroid hormone in tissue culture. *Nature* **197,** 1015–1016 (1963).
66. Schinke, T. *et al.* Impaired gastric acidification negatively affects calcium homeostasis and bone mass. *Nature Med.***15,** 674–681 (2009).
67. Gong, Y. *et al.* LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. *Cell* **107,** 513–523 (2001).
    **This paper identifies *LRP5* as the gene mutated in OPPG.**
68. Little, R. D. *et al.* A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait. *Am. J. Hum. Genet.* **70,** 11–19 (2002).
69. Boyden, L. M. *et al.* High bone density due to a mutation in LDL-receptor-related protein 5. *N. Engl. J. Med.* **346,** 1513–1521 (2002).
    **This study identifies a mutation in *LRP5* as causing HBM syndrome in humans.**
70. Kato, M. *et al.* *Cbfa*1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. *J. Cell Biol.* **157,** 303–314 (2002).
71. Tamai, K. *et al.* LDL-receptor-related proteins in Wnt signal transduction. *Nature* **407,** 530–535 (2000).
72. Hay, E. *et al.* Interaction between LRP5 and Frat1 mediates the activation of the Wnt canonical pathway. *J. Biol. Chem.* **280,** 13616–13623 (2005).
73. Tolwinski, N. S. *et al.* Wg/Wnt signal can be transmitted through Arrow/LRP5,6 and Axin independently of Zw3/Gsk3β activity. *Dev. Cell* **4,** 407–418 (2003).
74. Li, X. *et al.* Sclerostin binds to LRP5/6 and antagonizes canonical Wnt signaling. *J. Biol. Chem.* **280,** 19883–19887 (2005).
75. Ye, X. *et al.* Norrin, frizzled-4, and Lrp5 signaling in endothelial cells controls a genetic program for retinal vascularization. *Cell* **139,** 285–298 (2009).
76. Yadav, V. K. *et al.* Lrp5 controls bone formation by inhibiting serotonin synthesis in the duodenum. *Cell* **135,** 825–837 (2008).
77. Jackson, A. *et al.* Gene array analysis of Wnt-regulated genes in C3H10T1/2 cells. *Bone* **36,** 585–598 (2005).
78. Glass, D. A. *et al.* Canonical Wnt signaling in differentiated osteoblasts controls osteoclast differentiation. *Dev. Cell* **8,** 751–764 (2005).
79. Kramer, I. *et al.* Osteocyte Wnt/β-catenin signaling is required for normal bone homeostasis. *Mol. Cell. Biol.* **30,** 3071–3085 (2010).
80. Reya, T. & Clevers, H. Wnt signalling in stem cells and cancer. *Nature* **434,** 843–850 (2005).
81. Cui, Y. *et al.* Lrp5 functions in bone to regulate bone mass. *Nature Med.* **17,** 684–691 (2011).
82. Saarinen, A. *et al.* Low density lipoprotein receptor-related protein 5

(*LRP5*) mutations and osteoporosis, impaired glucose metabolism and hypercholesterolaemia. *Clin. Endocrinol.* **72,** 481–488 (2010).
83. Frost, M. *et al.* Patients with high-bone-mass phenotype owing to *Lrp5-T253I* mutation have low plasma levels of serotonin. *J. Bone Miner. Res.* **25,** 673–675 (2010).
84. Frost, M. *et al.* Levels of serotonin, sclerostin, bone turnover markers as well as bone density and microarchitecture in patients with high bone mass phenotype due to a mutation in Lrp5. *J. Bone Miner. Res.* **26,** 1721–1728 (2011).
85. Vilaca, T., Yamamoto, R. M., Carvalho, A. B. & Lazaretti-Castro, M. Neuroendocrine tumor associated with severe osteoporosis in a male patient. *Endocr Rev,* **32,** abstr. P3-123 (2011)
86. Modder, U. I. *et al.* Relation of serum serotonin levels to bone density and structural parameters in women. *J. Bone Miner. Res.* **25,** 415–422 (2010).
87. Shi, Z. C. *et al.* Modulation of peripheral serotonin levels by novel tryptophan hydroxylase inhibitors for the potential treatment of functional gastrointestinal disorders. *J. Med. Chem.* **51,** 3684–3687 (2008).
88. Liu, Q. *et al.* Discovery and characterization of novel tryptophan hydroxylase inhibitors that selectively inhibit serotonin synthesis in the gastrointestinal tract. *J. Pharmacol. Exp. Ther.* **325,** 47–55 (2008).
89. Inose, H. *et al.* Efficacy of serotonin inhibition in mouse models of bone loss. *J. Bone Miner. Res.* **26,** 2002–2011 (2011).
90. Yadav, V. K. *et al.* Pharmacological inhibition of gut-derived serotonin synthesis is a potential bone anabolic treatment for osteoporosis. *Nature Med.* **16,** 308–312 (2010).

**Acknowledgements** We thank T. Clemens, P. Ducy, M. Gershon, M. Kassem, S. Kousteni and members of the Karsenty laboratory for comments on the manuscript. We apologize to our colleagues whose work is not directly discussed and/or cited in this article because of space constraints. G.K. is supported by the National Institutes of Health, and M.F. by the Canadian Diabetes Association.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to G.K. (gk2172@columbia.edu).

nature podcast
science... wherever you are.
nature.com/nature/podcast
nature publishing group npg

nature

www.nature.com/nature/supplements/insights

REVIEW

doi:10.1038/nature10717

# The Amazon basin in transition

Eric A. Davidson[1], Alessandro C. de Araújo[2,3], Paulo Artaxo[4], Jennifer K. Balch[1,5], I. Foster Brown[1,6], Mercedes M. C. Bustamante[7], Michael T. Coe[1], Ruth S. DeFries[8], Michael Keller[9,10], Marcos Longo[11], J. William Munger[11], Wilfrid Schroeder[12], Britaldo S. Soares-Filho[13], Carlos M. Souza Jr[14] & Steven C. Wofsy[11]

**Agricultural expansion and climate variability have become important agents of disturbance in the Amazon basin. Recent studies have demonstrated considerable resilience of Amazonian forests to moderate annual drought, but they also show that interactions between deforestation, fire and drought potentially lead to losses of carbon storage and changes in regional precipitation patterns and river discharge. Although the basin-wide impacts of land use and drought may not yet surpass the magnitude of natural variability of hydrologic and biogeochemical cycles, there are some signs of a transition to a disturbance-dominated regime. These signs include changing energy and water cycles in the southern and eastern portions of the Amazon basin.**

Humans have been part of the vast forest–river system of the Amazon basin for many thousands of years, but expansion and intensification of agriculture, logging and urban footprints during the past few decades have been unprecedented. The human population of the Brazilian Amazon region increased from 6 million in 1960 to 25 million in 2010, and the forest cover for this region has declined to about 80% of its original area[1]. Efforts to curb deforestation have led to a steep decline in forest clearing in the Brazilian Amazon, from nearly 28,000 $\mathrm{km^2\,yr^{-1}}$ in 2004 to less than 7,000 $\mathrm{km^2\,yr^{-1}}$ in 2011[1]. However, this progress remains fragile. The river system produces about 20% of the world's freshwater discharge[2], and the forest biomass holds about 100 billion tonnes of carbon (C; refs 3, 4), which is equivalent to more than 10 years' worth of global fossil-fuel emissions. Maintaining the biotic integrity of the biome and the ecosystem services it provides to local, regional and global communities will require improved understanding of the vulnerability and resilience of Amazonian ecosystems in the face of change.

Here we provide a framework for understanding the linkages between natural variability, drivers of change, responses and feedbacks in the Amazon basin (Fig. 1). Although the basin-wide carbon balance remains uncertain, evidence is emerging for a directional change from a possible sink towards a possible source. Where deforestation is widespread at local and regional scales, the dry season duration is lengthening and wet season discharge is increasing. We show that the forest is resilient to considerable natural climatic variation, but global and regional climate change forcings interact with land-use change, logging and fire in complex ways, generally leading to forest ecosystems that are increasingly vulnerable to degradation.

## Natural and anthropogenic climatic variation

Changes in Amazonian ecosystems must be viewed in the context of the natural variation in climate[5,6] and soils[7] across the region, as well as natural cycles of climatic variation and extreme events. A climatic gradient spans the Amazon basin (Fig. 2), from the continuously rainy northwest to the wet/dry climate and long dry season of the southern and eastern regions, including the Cerrado (woodland/savannah) in the southeast. This climatic gradient is largely coincident with a gradient in land-use change, with more conversion to agriculture in the drier eastern and southern regions, indicating the interconnectedness of biophysical and socio-economic processes.

The El Niño/Southern Oscillation (ENSO) profoundly affects rainfall in the Amazon basin[5], especially the eastern portion; there is decreased flow of the Amazon River and some of its major tributaries during El Niño years, and increased flow and increased flooding during La Niña years[6]. The ENSO effect is superimposed over a 28-year cycle of variation in precipitation[5,6] such that the biggest floods occur when La Niña coincides with the wet phase in the 28-year cycle; this coincidence last occurred in the mid-1970s (Fig. 3). The worst droughts occur when El Niño coincides with the dry phase of the longer-term cycle, such as the 1992 drought. The North Atlantic Oscillation (NAO) also affects the region, contributing to, for example, the 2005 drought, which resulted in the lowest river levels recorded until then in southern and western tributaries[8]. Although much has been learned about extreme events and decadal-scale cycles, no discernable long-term trend has yet been identified in the total discharge of the Amazon River[9].

## Forests are resistant to seasonal droughts

The ability of roots to access deep soil water[10] and to redistribute it[11] helps to maintain evergreen canopies during dry seasons, demonstrating the adaptation of Amazon forest species to seasonal drought. The combination of access to deep soil water and less cloudiness permits continued plant photosynthesis throughout most of the dry season[12]. However, transitional forests and Cerrado ecosystems, where mean annual precipitation is less than 1,700 mm and the dry season lasts for ≥4 months, show clear evidence of dry season declines in evapotranspiration and therefore potential water stress[13]. Many tree species in the Amazon and Cerrado produce a flush of new green leaves near the end of every dry season, which is often detected in satellite images as an

---

[1] The Woods Hole Research Center, 149 Woods Hole Road, Falmouth, Massachusetts 02540-1644, USA. [2] Embrapa Amazônia Oriental, Travessa Dr. Enéas Pinheiro, s/n, Marco, Caixa Postal 48, Belém, Pará 66095-100, Brazil. [3] Instituto Nacional de Pesquisas da Amazonia (INPA), Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA), Avenida André Araújo, 2936, Manaus, Amazonas, 69060-001, Brazil. [4] Instituto de Física, Universidade de São Paulo, Rua do Matão, Travessa R, 187, São Paulo, SP 05508-090, Brazil. [5] National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, California 93101, USA. [6] Universidade Federal do Acre, Mestrado em Ecologia e Manejo de Recursos Naturais, Parque Zoobotânico, Distrito Industrial, Rio Branco, AC 69915-900, Brazil. [7] Universidade de Brasília, Instituto de Ciências Biológicas, Departamento de Ecologia, Campus Universitário Darcy Ribeiro, Asa Norte, Brasília, Distrito Federal 70910-900, Brazil. [8] Columbia University, Department of Ecology, Evolution, and Environmental Biology, 1200 Amsterdam Avenue, New York, New York 10027, USA. [9] USDA Forest Service, International Institute of Tropical Forestry, Jardín Botánico Sur, 1201 Calle Ceiba, San Juan, Puerto Rico 00926-1119, USA. [10] Embrapa Monitoramento por Satélite, Avenida Soldado Passarinho, 303, Fazenda Chapadão, Campinas, São Paulo, Brazil. [11] Harvard University, School of Engineering and Applied Sciences, Department of Earth and Planetary Sciences, 20 Oxford Street, Cambridge, Massachusetts 02138, USA. [12] University of Maryland, Earth System Science Interdisciplinary Center, 5825 University Research Court Suite 4001, College Park, Maryland 20740, USA. [13] Universidade Federal de Minas Gerais, Centro de Sensoriamento Remoto, Avenida Antônio Carlos 6627, Belo Horizonte, Minas Gerais 31270-900, Brazil. [14] Imazon, Centro de Geotecnologia do Imazon, Rua Domingos Marreiros 2020, Belém, Pará 66060-160, Brazil.

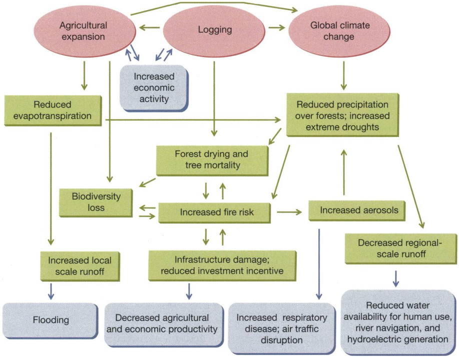

Agricultural expansion
Logging
Global climate change
Increased economic activity
Reduced evapotranspiration
Reduced precipitation over forests; increased extreme droughts
Forest drying and tree mortality
Biodiversity loss
Increased fire risk
Increased aerosols
Decreased regional-scale runoff
Increased local scale runoff
Infrastructure damage; reduced investment incentive
Flooding
Decreased agricultural and economic productivity
Increased respiratory disease; air traffic disruption
Reduced water availability for human use, river navigation, and hydroelectric generation

**Figure 1 | Interactions between global climate, land use, fire, hydrology, ecology and human dimensions.** Forcing factors are indicated with red ovals; processes addressed in this Review are indicated by green boxes and arrows; and consequences for human society are indicated by blue boxes with rounded corners.

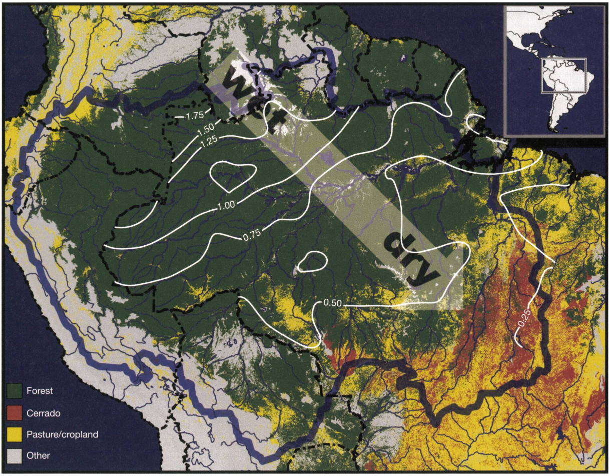

wet
dry
1.75
1.50
1.25
1.00
0.75
0.50
0.25
Forest
Cerrado
Pasture/cropland
Other

**Figure 2 | Climatic gradient across the Amazon basin.** Main figure, the hydrologic Amazon basin is demarcated by a thick blue line; isopleths of mean daily precipitation during the three driest months of the year[97] (in mm; white lines) are overlain onto four land-cover classes[98,99] (key at bottom left). These isopleths are presented only for areas within Brazil, because of lack of adequate data elsewhere. The arrow emphasizes the trend from continuously wet conditions in the northwest to long and pronounced dry seasons in the southeast, which includes Cerrado (savannah/woodland) vegetation. National boundaries are demarcated by broken black lines. Inset, map showing area of main figure (boxed).

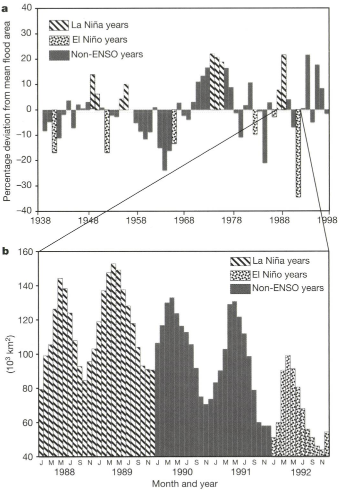


**Figure 3 | Decadal and seasonal variation in flood area.** The long-term record of simulated interannual variation in percentage deviation from mean flood area (**a**) illustrates how the ENSO events are superimposed over a 28-year cycle of high- and low-precipitation phases[100]. Recurrent seasonal patterns of flooded area ($10^3$ km$^2$) for five selected years (**b**) are superimposed over the interannual variation. Striped bars, La Niña years; stippled bars, El Niño years; solid bars, non-ENSO years.

increase in vegetation indices that involve ratios of red and near-infrared reflectance[14,15]. The relation between these satellite-based indices of seasonal greenness and ecosystem productivity remains an unresolved focus of debate in these studies, but in any case, this response represents a short-term phenomenon.

### Multi-year or extreme drought

Experimental manipulations and observations of permanent forest plots address responses to multi-year and extreme drought. Two long-term drought experiments have produced remarkably similar results, demonstrating that adaptation to seasonal drought can be overwhelmed by multi-year drought[16,17]. These studies demonstrated a physiological adaptation of the trees, which maintained a relatively constant water tension in the xylem (isohydry) in both wet and dry seasons; but this adaptation may eventually lead to mortality when roots are unable to extract enough soil water during multi-year droughts[18]. After diverting 35–50% of total rainfall for three years using below-canopy panels and gutters, plant-available soil moisture stores became depleted, wood production declined by about 30–60%, tree mortality nearly doubled, and live above-ground biomass decreased by about 18–25% (refs 16, 17). Mortality rates increased to nearly three times that in the control plot during years 4–7 of rainfall exclusion[17].

The severe 2005 drought in the southwestern Amazon—when dry season temperatures were 3–5 °C warmer than normal and rainfall over the Solimões River basin was only 33–65% of average values[8]—may have exceeded the adaptive capacity of many forest species. Analysis of 51 long-term monitoring plots across Amazonia showed that, relative to pre-2005 conditions, most forest plots subjected to increased water deficit in 2005 lost several tons of living tree biomass carbon per hectare, owing to a marginally non-significant decline in growth and a significant increase in mortality of trees[19]. A similarly severe but more extensive drought occurred again in 2010, affecting more than half of the basin and resulting in the lowest discharge ever recorded at Manaus[20,21]. Susceptibility to drought is likely to vary regionally, depending on the climate (total precipitation and its seasonal distribution) and soil water storage properties (texture and depth) to which the existing vegetation types (for example, Cerrado woodlands, tall-statured central Amazon forests, and transition forests) are physiologically adapted. Furthermore, there is evidence that certain taxa are more vulnerable to drought-induced mortality[17,22]. Despite this regional variability, the observations of natural droughts and the drought manipulation experiments indicate similar trends of mortality in response to dry season intensity[23].

## Land-use change and regional climate

Land use is changed to capture agricultural and forestry revenues, and results in trade-offs with multiple ecosystem services, such as C storage, climate regulation, hydrologic balance and biodiversity (Fig. 1).

### The drivers of deforestation

Road paving is one of the economic activities that stimulates deforestation[24]. Further clearing occurs along networks of 'unofficial roads' that result from the interacting interests of colonist farmers and loggers[25]; loggers minimize their costs by buying the right to log private lands. Although practices vary widely across the region, most small land holders (<200 ha) have kept more than 50% of their land in some combination of mature and secondary forest[26].

International and national demands for cattle and livestock feed are increasingly driving land-use change. Direct conversion of forest to cropland in 2003, mostly by large land holders, represented 23% of the deforestation in forest and Cerrado regions of the state of Mato Grosso[27]. Although cattle pasture remains the dominant use of cleared land, the growing importance of larger and faster conversion to cropland, mostly for soybean export, has defined a trend of forest loss in Amazonia since the early 2000s.

Although selective logging is not an immediate land-use change, it often leads to deforestation. From 1999 to 2003, the area annually logged in the Amazon basin was similar in magnitude to the area deforested[28]. Logged areas are accessible by logging roads and are likely to be cleared within only a few years after initial disturbance[29], and those that are not cleared have a high risk of burning[30]. On the other hand, reduced-impact logging has been demonstrated to be economically viable, while causing only modest and transient effects on carbon storage and water exchange[31]. Expansion of protected areas has also played an important role in reducing deforestation in the Brazilian Amazon (Fig. 4)[32].

### Deforestation alters the energy balance

Incoming air from the Atlantic Ocean provides about two-thirds of the moisture that forms precipitation over the Amazon basin[33]. The remainder is supplied through recycling of evapotranspiration, primarily driven by the deep-rooted Amazon trees.

A large number of observational and modelling studies have suggested that deforestation causes two main changes in the energy and water balance of the Amazon basin, as follows. First, partitioning of the net radiation that is absorbed by the land surface changes, with a decrease in the latent heat flux and an increase in the sensible heat flux, primarily because deforestation results in less vegetation being available to transpire water to the atmosphere. Second, replacing the dark rainforest with more reflective pasturelands or crops results in a decrease in solar radiation absorbed by the land surface. Reforestation can reverse these

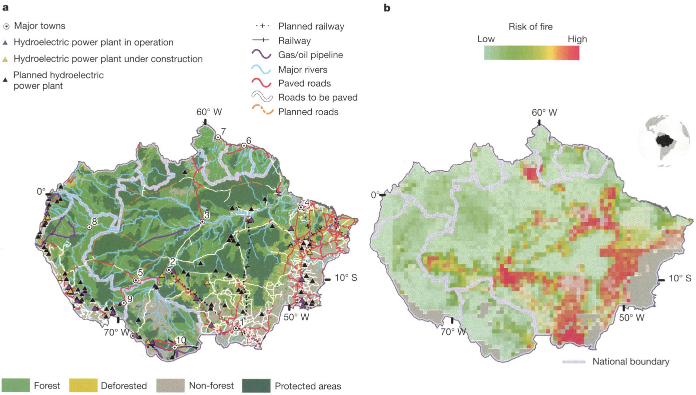


**Figure 4 | The Amazon basin today and future fire risks. a**, Protected areas and major planned infrastructure. **b**, The risk of fire by 2050[96] under business-as-usual deforestation and climate change scenarios[24]. The numbers in **a** indicate the following cities: 1, Cuiabá; 2, Porto Velho; 3, Manaus; 4, Belém; 5, Rio Branco; 6, Paramaribo; 7, Georgetown; 8, Iquitos; 9, Puerto Maldonado; 10, Santa Cruz de la Sierra.

trends. Within a few years of pasture abandonment, regrowing Amazonian forests establish rates of evapotranspiration and reflectivity that are close to those measured in mature forests, even though they have not yet recovered the biomass and species diversity of a mature forest[34,35].

Atmospheric convection and precipitation are driven by the fluxes of energy and water from the land surface. Where clearings for cattle pastures extend tens of kilometres outward from a road, the air above the deforested areas warms up more quickly and tends to rise and draw moist air from the surrounding forest, creating so-called 'vegetation breezes'. This decreases rainfall over the forest while increasing cloudiness, rainfall and thunderstorms over the pasture[36]. Heterogeneous deforestation at large scales (hundreds to thousands of km$^2$) leads to more complex circulation changes, with suppressed rainfall over core clearings, particularly at the beginning and the end of the wet season, and unchanged or increased rainfall over large remnant forest patches[37,38]. These changes also affect water and light availability, and the C uptake of the remaining forests, but those effects are not yet well quantified.

At deforestation scales greater than 10$^5$ km$^2$, numerical models consistently suggest that a significant decrease in basin-wide precipitation will occur[39] due to: (1) a decrease in the evapotranspiration from deforested regions and resultant downwind transport of water vapour; and (2) a decrease in net absorbed solar energy and a consequent general weakening of the continental-scale low-pressure system that drives precipitation over the basin.

## Deforestation, climate and river discharge

Taken alone, a decrease in regional precipitation would result in decreased discharge. However, the integrated response of a river system depends on the balance between precipitation and evapotranspiration effects (Fig. 1). Deforestation within a particular watershed would cause reduced evapotranspiration and increased discharge, but deforestation at the continental scale could cause reduced regional precipitation and a tendency towards decreased river discharge[39].

A large disturbance and a long data record are needed to detect unambiguously the effect of deforestation on the discharge from a large river, given the large interannual and decadal variation in precipitation. For most of the major tributaries of the Amazon River, the area deforested is not yet large enough to be able to attribute changes in discharge specifically to deforestation. Similarly, a temporal trend in sediment load could not be distinguished from highly variable interannual and seasonal variation for the Madeira River, which drains the southwestern Amazon basin[40]. However, for the Tocantins River[41] and Araguaia River[42] systems, which drain parts of the Cerrado and rainforest environments in the southeastern Amazon, the relative contributions of climate variability and deforestation have been teased apart. From 1955 to 1995, the area of pasture and cropland in the Tocantins basin increased from about 30% to 50% and annual river discharge increased by about 25%, but changes in precipitation were not statistically significant. Changes of the same magnitude have occurred in the Araguaia River since the 1970s, and sediment load increased by 28% with deforestation. In both rivers, discharge increased mostly during the wet season, when flooding risks are greatest. If deforestation approaches this magnitude in other tributaries, it is likely that land-use change will enhance flooding and sediment transport.

## Regional climate change

The IPCC fourth assessment climate change model runs show the highest probability of significant precipitation decrease predicted for southeastern Amazonia, where deforestation is greatest and where the climate and ecosystems transition from short-dry-season rainforest to long-dry-season savannah ecosystems[43,44]. Various global and regional climate modelling approaches have suggested that once deforestation exceeds about 40% of the entire Amazon basin, a 'tipping point' might be passed[45], whereby decreased energy and moisture released to the atmosphere from the largely deforested landscape would result in reduced convection and precipitation, and a shift in the forest–savannah boundary or large-scale dieback of rainforest.

A number of deficiencies in the structure and application of global climate models suggest that the uncertainty of these simulated outcomes and estimated tipping points is very high. A few examples include: (1) many of the models simulate too little precipitation in the Amazon owing to incomplete representation of the role of the Andes in continental circulation and large-scale convection over the core of the western Amazon, and also owing to coarse representation of the land surface with respect to small-scale meteorological processes[43]; (2) inter-annual variation of sea surface temperatures in the tropical Pacific and tropical north Atlantic Ocean are closely associated with extreme flood and drought events in the Amazon[5,8,20], but these teleconnections to the Amazon are not yet adequately represented in global climate models; and (3) the biophysical response of vegetation to increasing atmospheric $CO_2$, including effects on evapotranspiration, may be one of the largest unknowns for the future of the Amazon forests. The probability of simulated forest dieback due to decreased rainfall is greatly reduced when a strong $CO_2$ fertilization response is included in a vegetation model[44], but the scale of the actual impact of increasing $CO_2$ on photosynthetic efficiency remains a large source of uncertainty.

In summary, the changes in precipitation and discharge associated with deforestation already observed in the southern and eastern Amazon demonstrate a potential for significant vegetation shifts and further feedbacks to climate and discharge. Numerical models strongly suggest that potential future deforestation may also cause feedbacks to large-scale climate and vegetation distribution, but the models have deficiencies that prevent confident prediction of the magnitude or spatial distribution of deforestation that would lead to a significant region-wide decrease in precipitation—including whether a threshold, or tipping point, exists whereby the basin could slip into a dry, stable state. Focusing on a theoretical and difficult-to-define tipping point for the entire basin may divert the scientific community from the important large-scale regional changes that are already taking place, such as lengthening of the dry season[37,38] and increases in river discharge[41,42] in ecologically and agriculturally important transition zones of the eastern and southern flanks of the basin.

## Fire as cause and consequence of change

The probability of fire is clearly affected by climate and land use, the latter providing the majority of ignition sources today[46]. Fire also affects regional climate through a complex set of biophysical and socio-economic feedback processes (Fig. 1).

### Smoke changes cloud physics and rainfall

During the wet season, the air over most of the Amazon region is as pristine as air over the open ocean—only a few hundred aerosol particles per $cm^3$ of air[47]—inspiring the term 'green ocean'[48]. In stark contrast, burning for land clearing, pasture management and charcoal production, and escaped forest fires during the dry season, increase aerosols to more than 40,000 particles per $cm^3$ of air in some regions[47]. This smoke and haze affects the microphysical processes within clouds that determine how droplets are formed, making droplets too small to precipitate as rain, thus reducing local rainfall and increasing cloud lifetime[49]. The water vapour remaining in the atmosphere ascends to higher altitudes, where it invigorates thunderstorm formation and lightning strikes, but not necessarily rain. During the dry season, satellite-based measurements of aerosol optical depth were inversely correlated with precipitation[50]. In addition to locally smoke-inhibited rainfall, fires cause further plant stress due to ozone pollution[51] and thick haze that reduces light availability and photosynthesis[52]. Generally, plants are most productive with some scattered light at intermediate levels of aerosol thickness, but conditions during the biomass burning season often exceed this optimum[52].

### Drought increases fire susceptibility

The tall, dense tree canopy of central Amazonian forests creates a humid microclimate at ground level, which naturally protects the forest from fire[53]. However, several lines of evidence indicate that this natural resistance may be changing: (1) about 39,000 $km^2$ of Amazon forest burned during the El Niño drought of 1998[54], including intact, closed-canopy forests; (2) both logging and drought-induced tree mortality allow sunlight to penetrate clearings in the canopy, which dries out the forest floor, rendering it more flammable[53]; (3) after a forest is burned once, it is more likely to burn again, because a burned forest dries out more easily[30]; and (4) ignition sources have also increased owing to pasture management and charcoal making[46]. Although Brazil has made great strides in recent years to reduce rates of deforestation[1], the frequency of fire has not decreased[55], and prospects for continued forest degradation resulting from fires escaping nearby agricultural areas may be a growing risk in many regions.

### Fires alter forest characteristics

After fires sweep through Amazonian forests, tree mortality ranges from 8% to 64% of mature stems ($\geq$10 cm diameter at breast height)[56]. More frequent and/or more severe fires tend to increase tree and liana mortality[57]. Big trees are generally better adapted to surviving fire, but tend to be the first to suffer from drought[19,22,57]. Although surviving stems can benefit from the initial pulse of fire-released nutrients and reduced competition, fire-induced mortality reduces overall canopy cover, biomass, and species richness[57,58]. The decline in plant species diversity also reduces the abundance of fruits and invertebrates, thereby changing the food supply of birds and other animals[56]. Frequent fire could change the structure, composition and functioning of vegetation by selecting fire-adapted species and favouring more flammable species (for example, grasses), thus leading to a more savannah-like ecosystem[59].

### Multiple fires retard forest regrowth

Fire is used as a tool to help clear land for cattle pasture and to slow the invasion of woody shrubs, but pastures are often abandoned after a few years, when grass productivity declines and weeds can no longer be effectively controlled. Despite tremendous diversity in rates of regrowth among secondary forests from different regions of the basin, the rate of secondary forest regrowth following pasture abandonment was found to be negatively correlated with the number of fires that occurred while in the pasture phase[60]. Nitrogen (N) loss during burning alters the natural patterns of phosphorus limitation on highly weathered soils. In a study of secondary forests growing on abandoned pastures and croplands, several soil and foliage indicators of N limitation were strongest in the youngest forest stands and became less pronounced as the forests aged[61]. After decades of forest regrowth, the N cycle gradually recuperates, establishing a N-rich mature forest, but the rate of recuperation, as well as the rate of forest regrowth, depends, in part, on the legacy of previous land use and fire.

## Disturbance effects on greenhouse gases

Changes in greenhouse-gas emissions due to disturbance processes must be placed in the context of natural emissions. Amazonian forests and wetlands are significant natural sources of methane[62–64] and nitrous oxide[65,66] (Fig. 5). Unfortunately, a net carbon balance for the region remains elusive.

### Mature forests may be accumulating carbon

Repeated sampling of about 100 permanent plots in the RAINFOR network scattered across nearly all Amazonian countries indicates that mature Amazonian forests have been accumulating carbon at an estimated rate of $0.4\,Pg\,C\,yr^{-1}$ ($1\,Pg = 10^{15}\,g$; 95% confidence interval range of estimate, $0.29–0.57\,Pg\,C\,yr^{-1}$) in the decades before the 2005 drought[19]. The fastest growing trees are in the foothills of the Andes, where the soils are generally younger and more fertile[7], but where the trees are generally smaller and shorter-lived[3,4]. In contrast, the biggest and slowest growing trees occur in the oldest and more nutrient-poor soils of the lowland central and eastern parts of the basin[67]. The soils of mature forests on highly weathered Oxisols and Ultisols are unlikely to

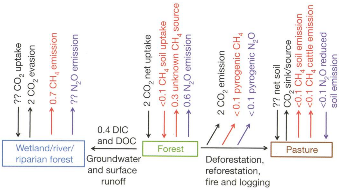

?? $CO_2$ uptake
2 $CO_2$ evasion
0.7 $CH_4$ emission
?? $N_2O$ emission
Wetland/river/ riparian forest
0.4 DIC and DOC
Groundwater and surface runoff
2 $CO_2$ net uptake
<0.1 $CH_4$ soil uptake
0.3 unknown $CH_4$ source
0.6 $N_2O$ emission
Forest
2 $CO_2$ emission
< 0.1 pyrogenic $CH_4$
< 0.1 pyrogenic $N_2O$
Deforestation, reforestation, fire and logging
?? net soil $CO_2$ sink/source
<0.1 $CH_4$ soil emission
<0.1 $CH_4$ cattle emission
<0.1 $N_2O$ reduced soil emission
Pasture

**Figure 5 | Estimates of Amazonian greenhouse-gas emissions.** Estimates of annual, basin-wide greenhouse-gas fluxes described in the text are presented together here, in a common currency of Pg $CO_2$-equivalents, using 100-year global warming potentials for $CO_2$ (black), $CH_4$ (red) and $N_2O$ (purple). Owing to large uncertainties, all values are rounded to one significant figure, and even these estimates remain subject to debate. Where no estimate is available, "??" is indicated. Note that dissolved inorganic carbon (DIC) and dissolved organic carbon (DOC) may be transported via groundwater and overland flow from upland forests to streamside (riparian) forests, and that $CO_2$ can be lost (evasion) from river water to the atmosphere.

be significant C sinks[68], although more study is needed on a wider diversity of soils.

The cause of observed biomass increases in mature Amazonian forests remains unknown. Plausible explanations include a rebound from previous human or natural disturbances[69] or a change in resources that limit plant productivity, such as atmospheric $CO_2$, soil nutrients, or light availability due to changes in radiation, climate and cloudiness[70]. The RAINFOR network is our best indicator of Amazonian above-ground biomass change, but the network is neither a systematic nor a randomized sampling of Amazonian vegetation. Because the network has relatively few small plots covering a vast region, the effects of large-scale natural disturbances over decadal and longer timescales may not be included in the sampling network, leading to an overestimate of a biomass increase[69,71]. Although this challenge to the RAINFOR conclusions has been rebutted[72,73], resolving the issue will require empirical data on the distribution of natural disturbances, which is still poorly known[71]. A recent analysis of satellite images and meteorological data showed that large disturbances (>5 ha) caused by windstorms are rare, with a return interval of about 40,000 years (ref. 74), suggesting that such disturbance effects may not be common enough to undermine extrapolations of carbon uptake rates from the RAINFOR network. However, more work on this topic is needed, including better estimates of the return intervals of smaller disturbances (<5 ha)[71].

Full C accounting should also include exports from forests to aquatic systems. The river water is supersaturated with dissolved $CO_2$, which is eventually released to the atmosphere at an estimated rate of about 0.5 Pg C $yr^{-1}$ (ref. 75). Estimates of the sources of this C remain poorly constrained—about two-thirds may come from leaf and wood detritus dropped into the river from flooded forests, with about one-third produced by aquatic plants (mats of grasses and other macrophytes) within the river, and a small fraction by algae[76]. Additional possible sources include particulates washed in with soil particles and dissolved organic and inorganic C in ground water[77,78]. We know very little about the C budget of flooded forests and riparian zone forests, which probably contribute significant terrestrially fixed C to streams and rivers[76].

Estimates of $CO_2$ fluxes based on year-round vertical profiling of atmospheric $CO_2$ concentrations by aircraft are available now only for part of the eastern Amazon. Fire emissions roughly cancel a modest biological sink during the dry season, so that a wet season source yields an annual net source of C to the atmosphere[79]. This result is consistent with ground-based estimates of slow growing trees and a concentration of land-use change in the eastern part of the basin. It remains to be seen if future aircraft measurements will corroborate the C sink inferred from scattered ground measurements in the more intact forests of the western part of the basin.

## Disturbing forests causes net C and N loss

The net effect of Amazonian deforestation and reforestation results in an annual net C source of 0.15–0.35 Pg C (ref. 80). Adding C emissions from fire and logging extends the range to an annual net release of 0.2–0.8 Pg C (ref. 80). The estimated mean annual C emission from deforestation and burning of Cerrado is 0.07 Pg C for 2003 to 2008[81]. These estimates are improving, in part because of advances in the technology for analysing satellite images[82] to combine spatially explicit deforestation rates with regionally specific estimates of forest-C stocks[83].

Pyrogenic $CH_4$ emissions from conversion of Amazonian and Cerrado native vegetation to pasture are about 1.0 and 0.4 Tg $CH_4$ $yr^{-1}$, respectively[81], but this does not include shifting cultivation or wildfire. Annual pyrogenic $N_2O$ emissions from conversion of Amazonian forest to pasture are about 0.01 Tg N as $N_2O$ (ref. 81), but this does not include shifting cultivation or wildfire.

When forests are replaced by cattle pastures, they can either gain[84] or lose[85] soil C. Losses are more common where soil C stocks are initially large, and gains are more common when management inputs (fertilizer, herd rotation, overgrazing avoidance) are greatest[86]. However, changes in soil C stocks are usually dwarfed by much larger losses in tree biomass. In contrast, the sparse and short-statured trees of the Cerrado have less above-ground biomass than an Amazonian forest, but the C stocks in roots and soil organic matter of the Cerrado (100 Mg C $ha^{-1}$ in the top 1 m of soil) can be 2–7 times higher than the above-ground stocks[87]. Well-managed cultivated pastures may provide enough C inputs to maintain soil C[88], but most pastures in the Cerrado region are in advanced stages of degradation, where C inputs are too low to sustain high soil C storage.

Amazonian upland forest soils annually take up about 1–3 Tg of $CH_4$, and pasture soils are probably a small net annual source of <0.1 Tg $CH_4$ (ref. 65). A significant net emission of $CH_4$ in upland forests has been measured, which might include termites or anaerobic respiration in waterlogged wood, soil, bromeliads, or moss patches, but the source remains unknown[64]. Enteric fermentation by cattle is estimated to emit 2.6 and 4.1 Tg $CH_4$ $yr^{-1}$ in Amazonian and Cerrado regions, respectively[81]. Continuing studies point to major hydroelectric reservoirs as an increasing source of methane[76]. On the basis of chamber flux measurements, upland Amazonian forest soils are estimated to emit 1.3 Tg $yr^{-1}$ of $N_2O$-N (ref. 65), which is about 15% of global non-anthropogenic emissions. Young cattle pastures have higher $N_2O$ emissions compared to forests, but old pastures have lower emissions, so the net effect of deforestation has been a small annual decrease of <0.1 Tg $N_2O$-N (ref. 65).

Although secondary forests may be significant carbon sinks in other parts of the world[89], they currently contribute little to the net C balance of the Amazon basin, because they are frequently re-cut before they grow large enough to store much C[90]. Indeed, the area of secondary forests is declining where agriculture continues to expand and intensify, leading to continued loss of biomass-C from those regions[91]. Agroforestry and other alternatives to slash-and-burn agriculture for smallholders have not been widely adopted, but the potential for significant C sequestration per hectare and the techniques of nutrient management in these systems have been demonstrated[92,93].

## Emerging evidence for a transition

Are impacts of land-use and climate change in the Amazon basin surpassing the natural variability of climate, greenhouse-gas emissions, and cycles of carbon, nitrogen and water? Thanks to increased research in this area, including the Large-scale Biosphere-Atmosphere (LBA) experiment in Amazonia (see the accompanying World View in this issue for a description of the LBA project), we can answer this question for some, but not all consequences of land-use and climate change. For greenhouse gases, the answer is probably 'not yet' with respect to $CH_4$

and $N_2O$, because they remain dominated by large emissions from undisturbed wetlands and soils, respectively, but the answer for $CO_2$ is more complex (Fig. 5). Although a C budget for the basin remains uncertain, deforestation has moved the net basin-wide budget away from a possible late-twentieth-century net C sink and towards a net source. This directional change is consistent with recent results of inverse modelling based on the TransCom3 network of $CO_2$ measurements, which reports a shift from a sink in the 1980s to a source in the 2000s for the tropical Americas[94]. Much of the Amazon forest is resilient to seasonal and moderate drought, but this resilience can and has been exceeded with experimental and natural severe droughts, indicating a risk of C loss if drought increases with climate change. The forest is also resilient to initial disturbances, but repeated or prolonged disturbance changes forest structure and nutrient dynamics, potentially leading to a long-term change in vegetation composition and C loss. A combination of regional net flux estimates based on aircraft campaign measurements with ground-based studies that elucidate process-level understanding is needed to narrow uncertainties.

With respect to energy and water cycles, at least two of the large river basins on the southeastern flanks of the Amazon forest that also drain the more heavily deforested Cerrado region—the Tocantins and Araguaia basins—have experienced increases in wet season discharge and sediment load. Evidence for changes in temporal and spatial patterns of precipitation, such as extended length of the dry season, is emerging at local and regional scales. We cannot yet answer the questions of whether total precipitation has changed or whether recent severe droughts and other extreme events are clear indicators of patterns expected to persist. Narrowing uncertainties about the effects of deforestation on regional precipitation, temperature and fire risk will require combining realistic spatial patterns of deforestation and degradation with improved mesoscale circulation models of climate.

The emerging evidence of a system in biophysical transition highlights the need for improved understanding of the trade-offs between land cover, carbon stocks, water resources, habitat conservation, human health and economic development in future scenarios of climate and land-use change[24,32,95,96]. Brazil is poised to become one of the few countries to achieve the transition to a major economic power without destroying most of its forests. However, continued improvements in scientific and technological capacity and human resources will be required in the Amazon region to guide and manage both biophysical and socio-economic transitions.

1. Instituto Nacional de Pesquisas Espaciais & National Institute for Space Research *Projeto Prodes Monitoramento da Florsta Amazonica Brasileira por Satélite Prodes* ⟨http://www.obt.inpe.br/prodes/⟩ (2011).
2. Salati, E. & Vose, R. Amazon basin: a system in equilibrium. *Science* **225,** 129–138 (1984).
   **One of the first presentations of the Amazon basin from a systems perspective.**
3. Malhi, Y. *et al.* The regional variation of aboveground live biomass in old-growth Amazonian forests. *Glob. Change Biol.* **12,** 1107–1138 (2006).
4. Saatchi, S. S., Houghton, R. A., Dos Santos Alvara, R. C., Soares, J. V. & Yu, Y. Distribution of aboveground live biomass in the Amazon basin. *Glob. Change Biol.* **13,** 816–837 (2007).
   **Estimates of regional variation and patterns in forest biomass are presented based on a remote sensing approach.**
5. Marengo, J. A. Interdecadal variability and trends of rainfall across the Amazon basin. *Theor. Appl. Climatol.* **78,** 79–96 (2004).
6. Coe, M. T., Costa, M. H., Botta, A. & Birkett, C. Long-term simulations of discharge and floods in the Amazon basin. *J. Geophys. Res.* **107,** 8044, http://dx.doi.org/10.1029/2001JD000740 (2002).
7. Quesada, C. A. *et al.* Variations in chemical and physical properties of Amazon forest soils in relation to their genesis. *Biogeosciences* **7,** 1515–1541 (2010).
8. Marengo, J. A., Nobre, C. A., Tomasella, J., Cardosa, M. F. & Oyama, M. D. Hydro-climate and ecological behaviour of the drought of Amazonia in 2005. *Phil. Trans. R. Soc. B* **363,** 1773–1778 (2008).
9. Marengo, J. in *Tropical Rainforest Responses to Climatic Change* (eds Bush, M. B. & Flenley, J. R.) 236–268 (Springer Praxis Books, 2007).
10. Nepstad, D. C. *et al.* The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures. *Nature* **372,** 666–669 (1994).
    **First demonstration of the importance of deep rooting for survival of eastern Amazonian trees.**
11. Oliveira, R. S. *et al.* Deep root function in soil water dynamics in cerrado savannas of central Brazil. *Funct. Ecol.* **19,** 574–581 (2005).
12. Saleska, S. R. *et al.* Carbon in Amazon forests: unexpected seasonal fluxes and disturbance-induced losses. *Science* **302,** 1554–1557 (2003).
13. da Rocha, H. R. *et al.* Patterns of water and heat flux across a biome gradient from tropical forest to savanna in Brazil. *J. Geophys. Res.* **114,** G00B12, http://dx.doi.org/10.1029/2007JG000640 (2009).
14. Brando, P., Goetz, S., Baccini, A., Nepstad, D. & Beck, P. Seasonal and interannual variability of climate and vegetation indices across the Amazon. *Proc. Natl Acad. Sci. USA* **107,** 14685–14690 (2010).
15. Huete, A. *et al.* Amazon rainforests green-up with sunlight in dry season. *Geophys. Res. Lett.* **33,** L06405, http://dx.doi.org/10.1029/02005GL025583 (2006).
16. Brando, P. M. *et al.* Drought effects on litterfall, wood production, and belowground carbon cycling in an Amazon forest: results of a throughfall reduction experiment. *Phil. Trans. R. Soc. B* **363,** 1839–1848 (2008).
17. da Costa, A. C. L. *et al.* Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytol.* **187,** 579–591 (2010).
18. Fisher, R. A., Williams, M., Lobo do Vale, R., Costa, A. & Meir, P. Evidence from Amazonian forests is consistent with isohydric control of leaf water potential. *Plant Cell Environ.* **29,** 151–165 (2006).
19. Phillips, O. L. *et al.* Drought sensitivity of the Amazon rainforest. *Science* **323,** 1344–1347 (2009).
20. Lewis, S. L., Brando, P. M., Phillips, O. L., van der Heijden, G. M. F. & Nepstad, D. C. The 2010 Amazon drought. *Science* **331,** 554 (2011).
21. Xu, L. *et al.* Widespread decline in greenness of Amazonian vegetation due to the 2010 drought. *Geophys. Res. Lett.* **38,** L07402, http://dx.doi.org/10.1029/2011GL046824 (2011).
22. Nepstad, D. C., Tohver, I. M., Ray, D., Moutinho, P. & Cardinot, G. Mortality of large trees and lianas following experimental drought in an Amazon forest. *Ecology* **88,** 2259–2269 (2007).
23. Phillips, O. L. *et al.* Drought–mortality relationships for tropical forests. *New Phytol.* **187,** 631–646 (2010).
24. Soares-Filho, B. S. *et al.* Modelling conservation in the Amazon basin. *Nature* **440,** 520–523 (2006).
    **Landmark presentation of scenarios of development and conservation policies in a spatially explicit simulation model.**
25. Arima, E. Y., Walker, R. T., Perz, S. G. & Caldas, M. M. Loggers and forest fragmentation: behavioral models of road building in the Amazon basin. *Ann. Assoc. Am. Geogr.* **95,** 525–541 (2005).
26. Brondízio, E. S. *et al.* in *Amazonia and Global Change* (eds Keller, M., Bustamante, M., Gash, J. & Dias, P. S.) 117–143 (American Geophysical Union, 2009).
27. Morton, D. C. *et al.* Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proc. Natl Acad. Sci. USA* **103,** 14637–14641 (2006).
28. Asner, G. P. *et al.* Selective logging in the Brazilian Amazon. *Science* **310,** 480–482 (2005).
29. Asner, G. P. *et al.* Condition and fate of logged forests in the Brazilian Amazon. *Proc. Natl Acad. Sci. USA* **103,** 12947–12950 (2006).
30. Nepstad, D. C. *et al.* Road paving, fire regime feedbacks, and the future of Amazon forests. *For. Ecol. Mgmt* **154,** 395–407 (2001).
31. Miller, S. D. *et al.* Reduced impact logging minimally alters tropical rainforest carbon and energy exchange. *Proc. Natl Acad. Sci. USA* http://dx.doi.org/10.1073/pnas.1105068108 **108,** 19431–19435 (2011).
32. Soares-Filho, B. S. *et al.* Role of the Brazilian Amazon protected areas in climate change mitigation. *Proc. Natl Acad. Sci. USA* **107,** 10821–10826 (2010).
33. Costa, M. H. & Foley, J. A. Trends in the hydrologic cycle of the Amazon basin. *J. Geophys. Res.* **104,** 14189–14198 (1999).
34. Hölscher, D., Sá, T. D. A., Bastos, T. X., Denich, M. & Fölster, H. Evaporation from young secondary vegetation in eastern Amazonia. *J. Hydrol.* **193,** 293–305 (1997).
35. Vieira, I. C. G. *et al.* Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sens. Environ.* **87,** 470–481 (2003).
36. Avissar, R. & Schmidt, T. An evaluation of the scale at which ground-surface heat flux patchiness affects the convective boundary layer using a large-eddy simulation model. *J. Atmos. Sci.* **55,** 2666–2689 (1998).
37. Butt, N., Oliveira, P. A. & Costa, M. H. Evidence that deforestation affects the onset of the rainy season in Rondonia, Brazil. *J. Geophys. Res.* **116,** D11120, http://dx.doi.org/10.1029/2010JD015174 (2011).
38. Knox, R., Bisht, G., Wang, J. & Bras, R. L. Precipitation variability over the forest to non-forest transition in southwestern Amazonia. *J. Clim.* **24,** 2368–2377 (2011).
39. Coe, M. T., Costa, M. H. & Soares-Filho, B. S. The Influence of historical and potential future deforestation on the stream flow of the Amazon River — land surface processes and atmospheric feedbacks. *J. Hydrol.* **369,** 165–174 (2009).
40. Leite, N. K. *et al.* Intra and interannual variability in the Madeira River water chemistry and sediment load. *Biogeochemistry* **105,** 37–51 (2011).
41. Costa, M. H., Botta, A. & Cardille, J. A. Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia. *J. Hydrol.* **283,** 206–217 (2003).
42. Coe, M. T., Latrubesse, E. M., Ferreira, M. E. & Amsler, M. L. The effects of deforestation and climate variability on the streamflow of the Araguaia River, Brazil. *Biogeochemistry* **105,** 119–131 (2011).
43. Malhi, Y. *et al.* Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proc. Natl Acad. Sci.* **106,** 20610–20615 (2009).
    **A review of climate model predictions for the Amazon basin.**
44. Rammig, A. *et al.* Estimating the risk of Amazonian forest dieback. *New Phytol.* **187,** 694–706 (2010).
45. Nobre, C. A. & Simone Borma, L. Tipping points' for the Amazon forest. *Curr. Opin. Environ. Sust.* **1,** 28–36 (2009).

46. Alencar, A., Solórzano, L. & Nepstad, D. Modeling forest understory fire in an eastern Amazonian landscape. *Ecol. Appl.* **14,** S139–S149 (2004).
47. Artaxo, P. *et al.* Physical and chemical properties of aerosols in the wet and dry season in Rondônia, Amazonia. *J. Geophys. Res.* **107** (D20), 8081–8095 (2002).
48. Williams, E. *et al.* Contrasting convective regimes over the Amazon: implications for cloud electrification. *J. Geophys. Res.* **107** (D20), 8082–8093 (2002).
49. Andreae, M. O. *et al.* Smoking rain clouds over the Amazon. *Science* **303,** 1337–1342 (2004).
**A review of understanding of how smoke from biomass burning affects local and regional climate.**
50. Bevan, S. L., North, P. R. J., Grey, W. M. F., Los, S. O. & Plummer, S. E. Impact of atmospheric aerosol from biomass burning on Amazon dry-season drought. *J. Geophys. Res.* **114,** D09204, http://dx.doi.org/10.1029/2008JD011112 (2009).
51. Longo, K. M. *et al.* Correlation between smoke and tropospheric ozone concentration in Cuiabá during SCAR-B. *J. Geophys. Res.* **104** (D10), 12113–12129 (1999).
52. Oliveira, P. H. F. *et al.* The effects of biomass burning aerosols and clouds on the $CO_2$ flux in Amazonia. *Tellus B* **59,** 338–349 (2007).
53. Ray, D., Nepstad, D. & Moutinho, P. Micrometeorological and canopy controls of flammability in mature and disturbed forests in an east-central Amazon landscape. *Ecol. Appl.* **15,** 1664–1678 (2005).
54. Alencar, A., Nepstad, D. C. & Vera Diaz, M. d. C. Forest understory fire in the Brazilian Amazon in ENSO and non-ENSO Years: area burned and committed carbon emissions. *Earth Interact.* **10,** 6, 1–17 (2006).
55. Aragão, L. E. O. & Shimabukuro, Y. E. The incidence of fire in Amazonian forests with implications for REDD. *Science* **328,** 1275–1278 (2010).
56. Barlow, J. & Peres, C. A. in *Emerging Threats to Tropical Forests* (eds Laurance, W. F. & Peres, C. A.) 225–240 (Univ. Chicago Press, 2006).
57. Balch, J. K. *et al.* Size, species, and fire characteristics predict tree and liana mortality from experimental burns in the Brazilian Amazon. *For. Ecol. Mgmt* **261,** 68–77 (2011).
58. Balch, J. D. *et al.* Negative fire feedback in a transitional forest of southeastern Amazonia. *Glob. Change Biol.* **14,** 2276–2287 (2008).
59. Nepstad, D. C., Stickler, C. M., Soares-Filho, B. & Merry, F. Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. *Phil. Trans. R. Soc. B* **363,** 1737–1746 (2008).
**Explores the mechanisms of how land use, fire and climate change interact.**
60. Zarin, D. J. *et al.* Legacy of fire slows carbon accumulation in Amazonian forest regrowth. *Front. Ecol. Environ.* **3,** 365–369 (2005).
61. Davidson, E. A. *et al.* Recuperation of nitrogen cycling in Amazonian forests following agricultural abandonment. *Nature* **447,** 995–998 (2007).
**Chronosequences of secondary forests were analysed to demonstrate that nitrogen limitation occurs in young Amazonian forests and then gradually declines during secondary succession.**
62. Melack, J. M. *et al.* Regionalization of methane emissions in the Amazon basin with microwave remote sensing. *Glob. Change Biol.* **10,** 530–544 (2004).
63. Miller, J. B. *et al.* Airborne measurements indicate large methane emissions from the eastern Amazon basin. *Geophys. Res. Lett.* **34,** L10809, http://dx.doi.org/10.1029/2006GL029213 (2007).
64. do Carmo, J. B., Keller, M., Dias, J. D., de Camargo, P. B. & Crill, P. A source of methane from upland forests in the Brazilian Amazon. *Geophys. Res. Lett.* **33,** 1–4 http://dx.doi.org/10.1029/2005GL025436 (2006).
65. Davidson, E. A. & Artaxo, P. Globally significant changes in biological processes of the Amazon Basin: results of the Large-scale Biosphere-Atmosphere Experiment. *Glob. Change Biol.* **10,** 519–529 (2004).
66. D'Amelio, M. T. S., Gatti, L. V., Miller, J. B. & Tans, P. Regional $N_2O$ fluxes in Amazonia derived from aircraft vertical profiles. *Atmos. Chem. Phys.* **9,** 8785–8797 (2009).
67. ter Steege, H. N. *et al.* Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443,** 444–447 (2006).
68. Telles, E. C. C. *et al.* Influence of soil texture on carbon dynamics and storage potential in tropical forest soils of Amazonia. *Glob. Biogeochem. Cycles* **17,** 1040, http://dx.doi.org/10.1029/2002GB001953 (2003).
69. Fisher, J. I., Hurtt, G. C., Thomas, R. Q. & Chambers, J. Q. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.* **11,** 554–563 (2008).
70. Nemani, R. R. *et al.* Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* **300,** 1560–1563 (2003).
71. Chambers, J. Q. *et al.* Lack of intermediate-scale disturbance data prevents robust extrapolation of plot-level tree mortality rates for old-growth tropical forests. *Ecol. Lett.* **12,** E22–E25 (2009).
72. Gloor, M. *et al.* Does the disturbance hypothesis explain the biomass increase in basin-wide Amazon forest plot data? *Glob. Change Biol.* **15,** 2418–2430 (2009).
73. Lloyd, J., Gloor, E. U. & Lewis, S. L. Are the dynamics of tropical forests dominated by large and rare disturbance events? *Ecol. Lett.* **12,** E19–E21 (2009).
74. Espírito-Santo, F. D. B. *et al.* Storm intensity and old growth forest disturbances in the Amazon region. *Geophys. Res. Lett.* **37,** L11403, http://dx.doi.org/10.1029/2010GL043146 (2010).
75. Richey, J. E., Melack, J. M., Aufdenkampe, A. K., Ballester, V. M. & Hess, L. L. Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric $CO_2$. *Nature* **416,** 617–620 (2002).
**Presents first calculations of potential loss of $CO_2$ to the atmosphere from the Amazon River and its main tributaries.**
76. Melack, J. M., Novo, E. M. L. M., Forsberg, B. R., Piedade, M. T. F. & Maurice, L. in *Amazonia and Global Change* (eds Keller, M. *et al.*) 525–542 (American Geophysical Union Books, 2009).
77. Davidson, E. A., Figueiredo, R. O., Markewitz, D. & Aufdenkampe, A. Dissolved $CO_2$ in small catchment streams of eastern Amazonia: a minor pathway of terrestrial carbon loss. *J. Geophys. Res.* **115,** G04005, http://dx.doi.org/10.1029/2009JG001202 (2010).
78. Johnson, M. S. *et al.* $CO_2$ efflux from Amazonian headwater streams represents a significant fate for deep soil respiration. *Geophys. Res. Lett.* **35,** L17401, http://dx.doi.org/10.1029/2008GL034619 (2008).
79. Gatti, L. V. *et al.* Vertical profiles of $CO_2$ above eastern Amazonia suggest a net carbon flux to the atmosphere and balanced biosphere between 2000 and 2009. *Tellus B* http://dx.doi.org/10.1111/j.1600-0889.2010.00484.x (published online, 6 July 2010).
80. Houghton, R. A., Gloor, M., Lloyd, J. & Potter, C. in *Amazonia and Global Change* (eds Keller, M. *et al.*) 409–428 (American Geophysical Union Books, 2009).
**The net effect of carbon loss due to deforestation and carbon accumulation from forest regrowth is estimated.**
81. Bustamante, M. M. C. *et al.* Estimating greenhouse gas emissions from cattle raising in Brazil. *Clim. Change* (submitted).
82. Morton, D. C. *et al.* Rapid assessment of annual deforestation in the Brazilian Amazon using MODIS data. *Earth Interact.* **9,** 1–22 (2005).
83. Fearnside, P. M. *et al.* Biomass and greenhouse-gas emissions from land-use change in Brazil's Amazonian "arc of deforestation": the states of Mato Grosso and Rondonia. *For. Ecol. Mgmt* **258,** 1968–1978 (2009).
84. Cerri, C. E. P. *et al.* Modelling changes in soil organic matter in Amazon forest to pasture conversion, using the Century model. *Glob. Change Biol.* **10,** 815–832 (2004).
85. Asner, G. P., Townsend, A. R., Bustamante, M. M. C., Nardoto, G. B. & Olander, L. P. Pasture degradation in the Central Amazon: linking changes in carbon and nutrient cycling with remote sensing. *Glob. Change Biol.* **10,** 844–862 (2004).
86. Neill, C. & Davidson, E. A. in *Global Climate Change and Tropical Ecosystems* (eds Lal, R., Kimble, J. M. & Stewart, B. A.) 197–211 (CRC Press, 2000).
87. Grace, J., San Jose, J., Meir, P., Miranda, H. S. & Montes, R. A. Productive and carbon fluxes of tropical savannas. *J. Biogeogr.* **33,** 387–400 (2006).
88. Santos, A. J. B. *et al.* High rates of net ecosystem carbon assimilation by Brachiara pasture in the Brazilian cerrado. *Glob. Change Biol.* **10,** 877–885 (2004).
89. Pan, Y. *et al.* A large and persistent carbon sink in the world's forests. *Science* **333,** 988–993 (2011).
90. Neeff, T., Lucas, R. M., Santos, J. d., Brondizio, E. S. & Freitas, C. C. Area and age of secondary forests in Brazilian Amazonia 1978–2002: an empirical estimate. *Ecosystems* **9,** 609–623 (2006).
91. Almeida, A. S. d., Stone, T. A., Vieira, I. C. G. & Davidson, E. A. Non-frontier deforestation in the eastern Amazon. *Earth Interact.* **14,** 1–15 (2010).
92. Luizão, F., Fearnside, P. M., Cerri, C. E. P. & Lehmann, J. in *Amazonia and Global Change* (eds Keller, M., Bustamante, M., Gash, J. & Dias, P. S.) 311–336 (American Geophysical Union, 2009).
93. Davidson, E. A. *et al.* An integrated greenhouse gas assessment of an alternative to slash-and-burn agriculture in eastern Amazonia. *Glob. Change Biol.* **14,** 998–1007 (2008).
94. Gurney, K. R. & Eckels, W. J. Regional trends in terrestrial carbon exchange and their seasonal signatures. *Tellus B* **63,** 328–339 (2011).
95. da Silva, R. R., Werth, R. D. & Avissar, R. Regional impacts of future land-cover changes on the Amazon Basin wet-season climate. *J. Clim.* **21,** 1153–1170 (2008).
96. Silvestrini, R. A. *et al.* Simulating fire regimes in the Amazon in response to climate change and deforestation. *Ecol. Appl.* **21,** 1573–1590 (2011).
97. Nepstad, D. C. *et al.* Amazon drought and its implications for forest flammability and tree growth: a basin-wide analysis. *Glob. Change Biol.* **10,** 704–717 (2004).
98. Eva, H. D. *et al.* A land cover map of South America. *Glob. Change Biol.* **10,** 731–744 (2004).
99. Sano, E. E., Rosa, R., Brito, J. L. & Ferreira, L. G. *Mapeamento de Cobertura Vegetal do Bioma Cerrado: Estratégias e Resultados* (Embrapa Cerrados, Planaltina, District Federal, Brazil, 2007).
100. Coe, M. T., Costa, M. H. & Howard, E. A. Simulating the surface waters of the Amazon River Basin: impacts of new river geomorphic and dynamic flow parameterizations. *Hydrol. Process.* **21,** 2542–2553 (2007).

**Acknowledgements** We thank the Brazilian Ministry of Science and Technology (MCT), the National Institute for Space Research (INPE) and the National Institute of Amazonian Research (INPA) for designing, leading and managing the LBA project. We also thank D. Wickland (NASA) for more than a decade of leadership and support for the LBA-Eco project component of LBA. We thank the LBA-Eco team members who contributed to discussions on an early draft of this manuscript at a workshop in Foz do Iguaçu in August 2010, and S. Saleska for comments on the manuscript. We thank P. Lefebvre and W. Kingerlee for assistance with figure and manuscript preparation. Development of this manuscript was supported by NASA grants NNX08AF63A and NNX11AF20G.

**Author Contributions** E.A.D. wrote an initial rough draft and edited the final draft of the paper. M.K. and M.T.C. contributed significant final edits. All of the other co-authors participated in an LBA-Eco team meeting at Foz do Iguaçu on August, 13, 2010, where this manuscript was designed, and either were co-leaders of breakout groups or made significant subsequent contributions of subsections of text or figures. All co-authors also provided edits throughout.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence should be addressed to E.A.D. (edavidson@whrc.org).

doi:10.1038/nature10733

# A novel retinoblastoma therapy from genomic and epigenetic analyses

Jinghui Zhang[1]*, Claudia A. Benavente[2]*, Justina McEvoy[2]*, Jacqueline Flores-Otero[2]*, Li Ding[3,4], Xiang Chen[1], Anatoly Ulyanov[1], Gang Wu[1], Matthew Wilson[5,6], Jianmin Wang[7], Rachel Brennan[2], Michael Rusch[1], Amity L. Manning[8], Jing Ma[9], John Easton[1], Sheila Shurtleff[9], Charles Mullighan[9], Stanley Pounds[10], Suraj Mukatira[7], Pankaj Gupta[7], Geoff Neale[7], David Zhao[11], Charles Lu[3], Robert S. Fulton[3,4], Lucinda L. Fulton[3,4], Xin Hong[3,4], David J. Dooling[3,4], Kerri Ochoa[3,4], Clayton Naeve[11], Nicholas J. Dyson[8], Elaine R. Mardis[3,4,12], Armita Bahrami[9], David Ellison[9], Richard K. Wilson[3,4,13], James R. Downing[9] & Michael A. Dyer[2,5,14]

**Retinoblastoma is an aggressive childhood cancer of the developing retina that is initiated by the biallelic loss of *RB1*. Tumours progress very quickly following *RB1* inactivation but the underlying mechanism is not known. Here we show that the retinoblastoma genome is stable, but that multiple cancer pathways can be epigenetically deregulated. To identify the mutations that cooperate with *RB1* loss, we performed whole-genome sequencing of retinoblastomas. The overall mutational rate was very low; *RB1* was the only known cancer gene mutated. We then evaluated the role of RB1 in genome stability and considered non-genetic mechanisms of cancer pathway deregulation. For example, the proto-oncogene *SYK* is upregulated in retinoblastoma and is required for tumour cell survival. Targeting SYK with a small-molecule inhibitor induced retinoblastoma tumour cell death *in vitro* and *in vivo*. Thus, retinoblastomas may develop quickly as a result of the epigenetic deregulation of key cancer pathways as a direct or indirect result of *RB1* loss.**

Retinoblastoma is a rare childhood cancer of the retina that can develop in a sporadic or a heritable form and is fatal if untreated. When the *RB1* gene was first cloned, it was found to undergo biallelic inactivation in virtually all retinoblastoma tumours[1]. Since then, hundreds of genetic lesions have been identified in human cancer. These genetic lesions can be grouped on the basis of the signalling pathways they affect that have direct or indirect mechanistic links to many of the common cellular properties, or hallmarks, of cancer. Thus, the rate of cancer progression is related to the kinetics of acquisition of multiple genetic lesions and/or epigenetic changes that ultimately lead to activation of growth-signalling pathways, evasion of cell death and senescence, acquisition of limitless replicative potential, sustained angiogenesis, and local tissue invasion and metastasis[2].

*RB1* inactivation confers limitless replicative potential to retinoblasts and is rate limiting for retinoblastoma tumorigenesis[3]. However, the mechanisms that enable retinoblastoma cells to acquire the additional hallmarks of cancer remain unknown. Evidence from molecular, cellular and cytogenetic studies suggest that RB1 is required for maintaining chromosomal stability[4,5], and that its loss leads to chromosome instability in cells maintained in culture. These data raise the possibility that *RB1* inactivation may underlie the rapid acquisition of cooperating mutations in key cancer pathways through chromosome instability. Alternatively, epigenetic changes may have a more dominant role in cooperating with the loss of *RB1* retinoblastoma tumorigenesis. RB1 has been implicated in regulating most major epigenetic processes, including microRNA regulation, DNA methylation, histone modification and ATP-dependent chromatin reorganization[6–10]. Thus, inactivation of RB1 in retinoblasts may lead to the rapid epigenetic deregulation of cancer genes that contribute to the essential cellular properties of retinoblastoma.

In this study, the St Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project characterized the genetic landscape of retinoblastoma. Whole-genome sequencing (WGS) of four retinoblastomas and their paired germline DNA samples showed no genetic lesions in known tumour suppressor genes or oncogenes, other than *RB1* and *MYCN*. More importantly, an orthotopic xenograft derived from one of the primary tumours showed no evidence of clonal variation or new coding-region mutations. This finding suggests that the retinoblastoma genome is more stable than previously believed.

Unlike the genetic landscape of retinoblastoma, the epigenetic profile shows profound changes relative to that observed in normal retinoblasts. One of the most striking results was the induction of the expression of the proto-oncogene spleen tyrosine kinase (*SYK*) in human retinoblastoma. SYK is required for tumour cell survival, and inhibition of SYK with a small-molecule inhibitor caused the degradation of MCL1 and caspase-mediated cell death in retinoblastoma cells in culture and *in vivo*. These findings highlight how comprehensive genetic and epigenetic analyses of tumours can be integrated, leading to the discovery of promising new therapeutic approaches and shedding light on the mechanisms underlying the rapid progression of retinoblastoma following *RB1* inactivation.

## Retinoblastoma whole-genome sequencing

We performed WGS analysis on four primary human retinoblastoma samples (Supplementary Information, section 1, and Supplementary Table 1) and on matched normal tissue. Local tumour cell invasion,

---

[1]Department of Computational Biology and Bioinformatics, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [2]Department of Developmental Neurobiology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [3]The Genome Institute, Washington University School of Medicine in St Louis, St Louis, Missouri 63108, USA. [4]Department of Genetics, Washington University School of Medicine in St Louis, St Louis, Missouri 63108, USA. [5]Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA. [6]Department of Surgery, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [7]Hartwell Center for Biotechnology & Bioinformatics, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [8]Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA. [9]Department of Pathology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [10]Department of Biostatistics, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [11]Department of Information Sciences, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [12]Siteman Cancer Center, Washington University School of Medicine in St Louis, St Louis, Missouri 63108, USA. [13]Department of Medicine, Washington University School of Medicine in St Louis, St Louis, Missouri 63108, USA. [14]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA.

*These authors contributed equally to this work.

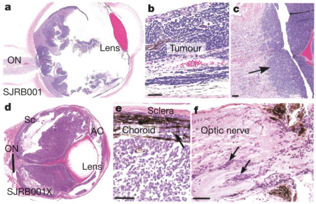


**Figure 1 | Characterization of retinoblastomas samples. a–c**, Representative retinoblastoma tumour section (SJRB001) stained with haematoxylin and eosin (H&E), showing choroidal and optic nerve invasion (arrow). **d–f**, H&E-stained section of the SJRB001X orthotopic xenograft with choroidal (**e**) and optic nerve (**f**) invasion (arrows). AC, anterior chamber; ON, optic nerve; Sc, sclera. Scale bars, 25 μm.

but not metastasis, was evident in each patient (Fig. 1a–c and Supplementary Fig. 1). We generated an orthotopic xenograft, SJRB001X, of the primary tumour SJRB001 by inoculating primary tumour cells into the vitreous humour of the eyes of immunocompromised mice (Supplementary Information, section 2). SJRB001X had molecular, genetic and histopathologic features similar to those of SJRB001 (Fig. 1d–f; Supplementary Figs 2–4; Supplementary Tables 2–4; and Supplementary Information, section 3).

Using a paired-end sequencing approach, we generated 1,040.9 gigabase pairs (Gb) of sequence data for the samples described; 956.8 Gb (92%) was successfully mapped to the NCBI 36.1 reference genome (Supplementary Information, section 4, and Supplementary Table 5). The average genome coverage was ×28.9, and the average exon coverage was ×23.8 with 98.4% of single nucleotide polymorphisms (SNPs) detected across all nine genomes showing concordance with their corresponding SNP array genotype calls at the same genomic positions (Supplementary Table 5). To provide additional sequence coverage, we performed transcriptome sequencing of all four primary tumours (Supplementary Information, section 5, and Supplementary Table 6).

We identified 668 validated somatic sequence mutations and 40 structural variations across the four cases (Table 1). These included 23 tier-1 mutations in genes, 35 tier-2 mutations in evolutionarily conserved regions of the genome (Supplementary Information, section 6), 309 tier-3 mutations in non-repetitive regions of the genome that are not part of tiers 1 and 2, and 301 tier-4 mutations in repetitive sequences in the genome (Table 1 and Supplementary Table 7). The average number of sequence mutations was 167 per case (range, 56–258), with only 3.25 mutations per case (range, 0–5) resulting in amino-acid changes (Table 1). The estimated mean mutation rate was $6.7 \times 10^{-8}$ per base (range, $1.03 \times 10^{-7}$–$2.17 \times 10^{-8}$), which is 15-fold less than that in adult tumours analysed by WGS, except for acute myeloid leukaemia[11]. The predominant changes were C→A and G→T transversions (Supplementary Fig. 5), which is consistent with the possibility that some of the transversions result from production of 8-oxoguanine during oxidative stress. Tumour SJRB002 had no somatic sequence variations that resulted in amino-acid changes; the only structural variations were the loss of heterozygosity (LOH) at the *RB1* locus on chromosome 13 and a gain of chromosome 6p. This suggests that very few genetic lesions are required for retinoblastoma progression after *RB1* inactivation.

## *RB1* inactivation in retinoblastoma

Both *RB1* alleles were inactivated in each sample (Supplementary Figs 6 and 7). Tumours SJRB002 and SJRB003 had mutations in *RB1* combined with copy-number-neutral LOH, and SJRB001 and SJRB004 had somatic sequence mutations combined with *RB1*-promoter hypermethylation (Supplementary Figs 6–8). Deep-sequence analysis of the germline sample from SJRB002 revealed that about 10% of reads contained the R445 nonsense mutation, suggesting the presence of germline chimaerism for the *RB1* mutation (Supplementary Fig. 6). Combining the WGS data with SNP array data of an additional 42 samples, we found that tumours from patients with lower regional nucleotide diversity were much less likely to undergo LOH at the *RB1* locus (Supplementary Tables 8 and 9 and Supplementary Information, section 7). These data show a significant association ($P = 8 \times 10^{-8}$, Fisher's exact test) between a germline genetic variation and mechanism of biallelic *RB1* inactivation in retinoblastoma.

## Recurrent lesions in retinoblastoma

To determine whether any of the 11 genes with somatic mutations that caused amino-acid changes or a frameshift in the coding region (Table 1; Fig. 2a, b; Supplementary Figs 9 and 10; Supplementary Information, section 8; and Supplementary Table 10) were recurrently mutated in retinoblastoma, we sequenced all exons from the 11 genes in our recurrent screening cohort of 42 retinoblastomas (Supplementary Information, section 4). Only *BCOR* was recurrently mutated in retinoblastoma (6 of 46, or 13%). Five of the samples had *BCOR* mutations that resulted in truncation of the encoded protein, and one sample had a focal gene deletion (Fig. 2c, Supplementary Table 11 and Supplementary Fig. 11).

We also used the WGS data to identify somatic structural variations including whole-chromosome gains and losses, focal deletions (DEL), insertions (INS), inversions (INV), intrachromosomal rearrangements (ITX), interchromosomal rearrangements (CTX) and regions of LOH (Fig. 2a–c; Supplementary Information, section 9; Supplementary Table 12; and Supplementary Fig. 12). The average

**Table 1 | Somatic mutations and structural alterations in retinoblastoma**

| Sample* | Tier 1† | Non-silent tier 1‡ | Genes | Tier 2‖ | Tier 3¶ | Tier 4# | Total | Mutation rate | Structural variations** |
|---|---|---|---|---|---|---|---|---|---|
| SJRB001 D, G | 7 | 4 | *RB1, CCNC, AGMO, RHBG* | 16 | 117 | 85 | 225 | $1.03 \times 10^{-7}$ | 4 |
| SJRB001 X, D§ | 0 | 0 | NA | 8 | 68 | 9 | 85 | $5.87 \times 10^{-8}$ | 4 |
| SJRB002 D, G | 1 | 0 | NA | 1 | 25 | 29 | 56 | $2.17 \times 10^{-8}$ | 0 |
| SJRB003 D, G | 7 | 4 | *RB1, HNMT, LHX8, STOML2* | 5 | 67 | 50 | 129 | $5.79 \times 10^{-8}$ | 24 |
| SJRB004 D, G | 8 | 5 | *RB1, CD300LG, SDK1, TXK, DMWD* | 13 | 100 | 137 | 258 | $8.63 \times 10^{-8}$ | 12 |

Background mutation rate was calculated on the basis of the ratio of tier-3 mutations to tier-3 bases covered at least tenfold in tumour and germline samples for each pair.

* D, diagnostic tumour sample; G, germline (blood DNA) sample; X, xenograft sample.

† Tier 1 mutations are found in genes and include exons, 5′ and 3′ untranscribed regions, and splice sites. Introns are not included.

‡ Non-silent tier-1 mutations change amino acids in genes.

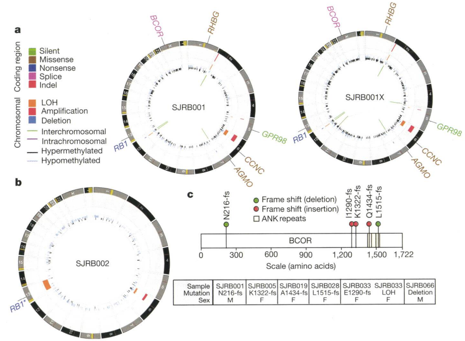
§ All of the somatic mutations in SJRB001 D, G were identified in SJRB001X. This row highlights the new mutations acquired in the xenograft relative to the primary tumour.

‖ Tier-2 mutations are found in regions of the genome that are conserved between humans and mice.

¶ Tier-3 mutations are found in regions of the genome that are not evolutionarily conserved.

# Tier-4 mutations are found in repetitive regions of the genome.

** Structural variations include focal amplifications and deletions, LOH, and interchromosomal and intrachromosomal translocations.



| Sample | SJRB001 | SJRB005 | SJRB019 | SJRB028 | SJRB033 | SJRB033 | SJRB066 |
|---|---|---|---|---|---|---|---|
| Mutation | N216-fs | K1322-fs | A1434-fs | L1515-fs | E1290-fs | LOH | Deletion |
| Sex | M | F | F | F | F | F | M |

**Figure 2 | Genomic profiles of SJRB001, SJRB002 and SJRB001X.** **a**, **b**, CIRCOS plots of genetic alterations in two retinoblastomas and the matched orthotopic xenograft. Loss of heterozygosity (orange), amplifications (red) and deletions (blue) are shown. Interchromosomal translocations (green lines) and intrachromosomal translocations (purple lines) are indicated. Sequence mutations in NCBI RefSeq genes included silent single nucleotide variants (SNVs; green), missense SNVs (brown), nonsense SNVs (dark blue), splice-site mutations (pink) and insertion/deletion mutations (indels, red). *Germline mosaic. **c**, *BCOR* mutations identified in the discovery and recurrency cohort.

number of structural variations was ten per case (range, 0–24) (Supplementary Table 1). Tumour SJRB001 had four structural variations (two DEL and ins INS) including a gain of a region of chromosome 2 spanning *MYCN* (Supplementary Table 12), and the only chromosomal lesion in SJRB002 was a gain of chromosome 6p, which occurs in about 40% of human retinoblastomas[12] (Fig. 2b). Only a few genomic regions were affected by the structural variations in SJRB003 and SJRB004 (Supplementary Figs 10, 13 and 14 and Supplementary Table 12).

## Orthotopic retinoblastoma xenograft

The genomic landscape of the orthotopic xenograft was remarkably similar to that of the primary tumour, despite continuous growth and multiple passages *in vivo* over nine months (Fig. 2 and Supplementary Table 1). All of the sequence mutations and structural variations detected in SJRB001 were retained in SJRB001X. Only 67 new SNVs and four structural variations were identified in the xenograft, and none affected annotated genes (Fig. 2 and Supplementary Table 1). Moreover, each mutation was identified at a subclonal level (range, 20–30%) and the mutant allele frequency for the lesions in tiers 1–4 in SJRB001 was retained in SJRB001X (Supplementary Fig. 15).

This result was surprising because several studies in mice and cell cultures have linked *RB1* inactivation to defects in chromosome segregation that result in aneuploidy[4,5,13–15] and chromosome instability[4]. We measured the distance between sister chromatids, the distance between kinetochores and the proportion of lagging chromatids in two *RB1*-deficient human retinoblastoma orthotopic xenografts[16] (SJRB001X and SJRB002X). Consistent with results from *RB1*-deficient retinal pigmented epithelium cells[4], the distances between sister chromatids and between kinetochores were increased, and there was evidence of lagging chromosomes (Supplementary Figs 16 and 17). However, less variation in ploidy was observed during spectral karyotype analysis of SJRB001X and SJRB002X, which was more consistent with the ploidy of wild-type cells (Fig. 3a, b and Supplementary Table 13). Moreover, copy number variations were much lower in our cohort of 46 retinoblastomas than in tumours with known genome instability such as ovarian cancer (Fig. 3c). Together, the cytogenetic data and WGS data suggest that the genome is stable and that newly acquired lesions do not provide a selective growth advantage and are thus probably passenger mutations (Supplementary Information, section 9).

## Identifying deregulated cancer pathways

There are many examples over the past several decades of epigenomic changes such as DNA methylation contributing to tumorigenesis[17–19]. Indeed, a recent study demonstrated changes in DNA methylation in Wilms' tumours[20], which tend, like retinoblastomas, to have stable genomes. To explore whether epigenetic deregulation of genes or pathways promotes tumorigenesis in retinoblastoma, we carried out an integrative analysis of chromatin immunoprecipitation (ChIP) data, DNA methylation data and gene expression data using order statistics. The SJRB001X sample was used for ChIP assay (Supplementary Figs 18–20), and primary tumour and xenograft samples were used for both DNA methylation and gene expression assays. In all three analyses, experimental results in retinoblastoma tumours were compared with those from human fetal retinae. In this comparison, a total of 104 genes, including 15 known cancer genes (Fig. 4a and Supplementary Tables 14 and 15), were found to have significant differences, indicating that several key cancer genes were epigenetically deregulated.

## SYK is a novel therapeutic target

*SYK* is the fifth most significant gene identified by the integrative analysis and the only upregulated kinase gene (Supplementary Table 15 and Fig. 4a). *SYK* is expressed throughout the haematopoietic system, regulates immunomodulatory signalling and has been implicated in several haematologic malignancies[21–24]. Small-molecule inhibitors of SYK have been developed to treat autoimmune disorders[25], and two of those agents, BAY 61-3606 and R406, have shown efficacy in preclinical leukaemia studies[26–28].

ChIP-on-chip analysis showed increased activating histone modifications (H3K4me3 and K3K9/14Ac) at *SYK*'s promoter, and the repressive histone marker (H3k9me3) was unchanged. Binding of

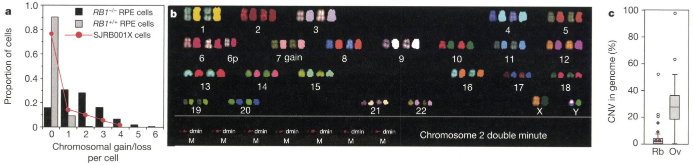


**Figure 3 | Analysis of aneuploidy and chromosome instability in retinoblastoma. a**, Chromosomal mis-segregation of SJRB001X cells after at least 21 rounds of cell division is plotted in red. RPE, retinal pigment epithelium. **b**, Representative spectral karyotype image of SJRB001X after the third passage in mice. **c**, Alterations in the 46 retinoblastoma (Rb) cases relative to 153 high-grade serous ovarian cancer (Ov) cases from The Cancer Genome Atlas. The median percentage of the genome involved in copy number variations (CNVs) was 1.5% for retinoblastoma and 27.7% for ovarian cancer. Red circles, samples used for WGS.

RNA polymerase II to the *SYK* promoter was also increased (Fig. 4b and Supplementary Tables 14 and 15). These ChIP-on-chip results were validated in independent samples by real-time PCR analysis with reverse transcription (Fig. 4c), and we confirmed the increase in *SYK* gene expression (Fig. 4d). SYK protein levels were higher in human retinoblastoma orthotopic xenografts and cell lines than in human fetal retinae (Fig. 4e). To determine whether SYK is expressed in primary human retinoblastomas, we performed immunohistochemistry on a retinoblastoma tissue microarray or whole-eye sections. In total, 100% (82 of 82) showed very strong expression (3+) of SYK in all tumour cells; SYK was not expressed in normal retina (Fig. 4f). SYK's kinase activity is regulated through autophosphorylation at the Tyr 525 and Tyr 526 residues within its catalytic domain. These sites were phosphorylated in retinoblastoma cells, and this phosphorylation was reversed by exposure to BAY 61-3606 or R406 (Fig. 4g and data not shown).

To determine whether *SYK* expression is required for retinoblastoma growth, survival or both, we generated a short hairpin RNA (shRNA)

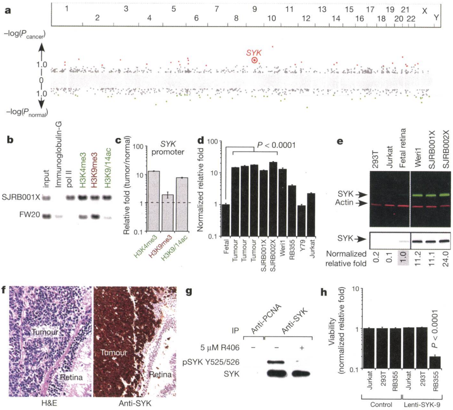


**Figure 4 | SYK is expressed in retinoblastoma and is required for survival. a**, Whole-genome view of the gene ranks based on integrating ChIP-on-chip, methylation and gene expression results. The *y* axis shows $-\log(P)$, where *P* is the *P* value of the *Q* statistic corrected for multiple testing. Significantly (false discovery rate, ≤10%) downregulated (green) or upregulated (red) genes are shown. **b**, **c**, ChIP validation of histone markers of the *SYK* promoter including quantification by quantitative PCR with TaqMan probes. In **c**, each bar shows mean ± s.d. of triplicate samples. **d**, *SYK* expression measured by quantitative PCR in fetal-week-20 retina (fetal), primary retinoblastoma (tumour), orthotopic xenografts (SJRB001X and SJRB002X) and cell lines. Each bar is mean ± s.d. of duplicate samples normalized to GPI1 expression. **e**, Immunoblot of SYK (green) and actin (red) in orthotopic xenografts, human fetal retina and representative cell lines; black and white representation of the SYK immunoblot is in the lower panel. **f**, H&E (purple) and anti-SYK (brown) immunohistochemistry of retinoblastoma tissue. **g**, Immunoprecipitation analysis of SYK and pSYK Y525/526 from Weri1 retinoblastoma cells. **h**, Viability was measured in triplicate cultures 72 h after infection of retinoblastoma cells with a lentivirus vector expressing either a control lentivirus or an shRNA against *SYK*. Scale bars, 10 μm.

against SYK and cloned it into the lentiviral vector Lenti-SYK-9. Lenti-SYK-9 efficiently knocked down *SYK* in retinoblastoma cell lines (Supplementary Fig. 21) and drastically increased apoptosis in retinoblastoma cells (Fig. 4h and Supplementary Fig. 21). Similar results were obtained *in vivo* using SJRB001X (data not shown). We used an empty lentiviral vector and a lentiviral vector encoding an SYK shRNA that less effectively reduced SYK expression (Lenti-SYK-6) as controls. Cell lines that do not express SYK (BJ, 293T and uninduced Jurkat cells) were used as controls and the Lenti-SYK-9 lentivirus had no effect on the growth or apoptosis of the control cells.

We exposed retinoblastoma cell lines that express high levels of *SYK* (Weri1 and RB355) to various concentrations of the SYK inhibitors BAY 61-3606 and R406 for 72 h and then measured cell viability. Jurkat (uninduced) and 293T cells were used as negative controls. Weri1 and RB355 cells were sensitive to both SYK inhibitors, but the Jurkat and 293T cells were unaffected (Fig. 5a). Transmission electron microscopy of retinoblastoma cells treated with the SYK inhibitors showed morphologic features consistent with cell death and mitochondrial defects (Supplementary Fig. 22); this was confirmed by scoring the proportion of activated caspase $3^+$ cells (Fig. 5b, c and Supplementary Fig. 22). Jurkat cells showed no increase in activated caspase $3^+$ cells after treatment with 10 μM R406 or BAY 61-3606 (data not shown).

The proportions of cells from each line that incorporated EdU were similar, suggesting that retinoblastoma's cell cycle is not affected by SYK inhibition (Fig. 5d, e and Supplementary Fig. 22). The effects of targeting SYK were partly rescued by the pan-caspase inhibitor Q-VD-OPH (Supplementary Fig. 22). Less MitoTracker Red accumulated in retinoblastoma cells exposed to the SYK inhibitors (Supplementary Fig. 22), yet treatment of Jurkat cells with either inhibitor had no effect on MitoTracker Red accumulation (data not shown). Inhibiting SYK in B-CLL cells reduced their autophosphorylation of SYK and destabilized MCL1[26,29], the only anti-apoptotic member of the BCL2 family that is upregulated in retinoblastoma (data not shown). Both inhibitors reduced autophosphorylation of SYK on Y525/526 (Fig. 4g) and reduced MCL1 expression coincident with apoptosis (Fig. 5f).

We tested the efficacy of BAY 61-3606 *in vivo* using our SJRB001X model[30]. The chemotherapy regimen consisted of a single subconjunctival dose of BAY 61-3606 on day 1 and daily doses of topotecan (TPT) on days 1–5 until either six courses (21 days per course) were administered or the tumour had progressed and surgical enucleation was required (Fig. 5g, h). BAY 61-3606 plus TPT significantly improved the outcome ($P = 0.003$; Fig. 5i), and its efficacy was correlated with an increase in activated caspase $3^+$ cells in the treated eyes (Fig. 5j). Previous studies using this model have shown that TPT combined with subconjunctival carboplatin had no effect on tumour response or outcome[30], so the improvement seen here can be attributed to targeting SYK. MCL1 expression was also reduced in the treated eyes, which is consistent with increased apoptosis and targeting of SYK *in vivo* (Fig. 5k). Together, these results suggest that SYK is a promising new target for treating retinoblastoma.

## Discussion

Genome sequencing reveals that retinoblastomas have a relatively stable genome. The mutational rate and number of structural variations per case that we assessed were among the lowest reported in human cancer to date. Moreover, in one example (SJRB002) the only non-silent mutation found was in *RB1*, and only two structural variations were detected. The minimal increase in passenger mutations in SJRB001X cells, despite prolonged passage, was also consistent with a relatively stable genome. These results are surprising because previous studies have shown that the functional inactivation of RB1 can cause

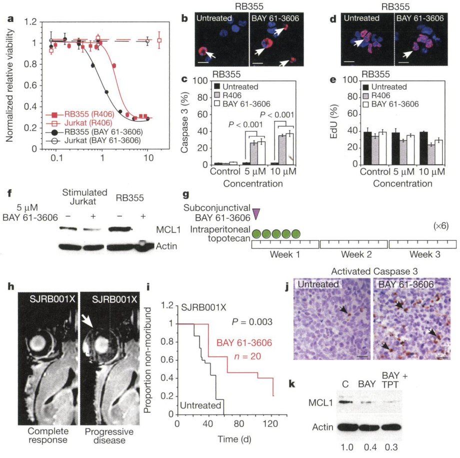


**Figure 5 | Retinoblastoma cells are sensitive to SYK inhibitors. a,** Dose response for SYK inhibitors R406 (red) and BAY 61-3606 (black) in RB355 retinoblastoma cells and a negative control (Jurkat). Each data point is mean ± s.d. of triplicate samples. **b–e,** Immunofluorescence of activated caspase 3 (**b**, **c**) or EdU (**d**, **e**) (red) before and after treatment of RB355 cells with R406 or BAY 61-3606. A total of 250 cells were scored in duplicate for each sample and each treatment condition to derive the mean and s.d. Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, blue). **f,** Treatment of stimulated Jurkat or RB355 cells with 5 μM BAY 61-3606 for 24 h reduced MCL1 expression. **g,** Schematic of the treatment schedule for mice with SJRB001X tumours. **h,** Representative magnetic resonance images of a mouse whose tumour responded after four courses of treatment with BAY 61-3606 (left) and another whose disease progressed during treatment (right). **i,** Survival curves show that treatment with BAY 61-3606 plus TPT improved outcome. **j,** Immunostaining for activated caspase 3 (arrows) in eyes either untreated or treated with BAY 61-3606. **k,** Immunoblot showing reduction in MCL1 after treatment with BAY 61-3606 (BAY) or BAY 61-3606 plus TPT. C, control. Scale bars, 5 μm (**b**, **d**); 10 μm (**j**).

genomic instability. Although retinoblastoma cells show the mitotic defects that can lead to errors in chromosome segregation, the ploidy of retinoblastoma cells is stable *in vitro* and *in vivo*[31,32]. We do not exclude the possibility that RB1 inactivation promotes change in other contexts, but our data show that genomic instability is not a hallmark of retinoblastoma and do not explain how retinoblastomas progress so rapidly.

We propose that epigenetic mechanisms contribute to retinoblastoma tumorigenesis. We identified several known oncogenes and tumour suppressor genes with histone modifications and altered DNA methylation that correlated with changes in gene expression. Our key discovery was that SYK is important in retinoblastoma. Retinal progenitor cells and retinal neurons express little or no SYK, and SYK has no known function in the developing visual system. Moreover, no recurrent genetic lesions in *SYK* were identified by WGS or SNP array analysis to suggest that this gene drives retinoblastoma tumorigenesis. Only by integrating epigenetic and gene expression analyses did we identify SYK as an important oncogene in retinoblastoma. This is important not only for expanding our understanding of the biology of retinoblastoma but also for advancing immediate therapeutic options that were not previously considered, such as the use of BAY 61-3606 or R406. This study highlights the value of integrating WGS analyses of the genetic and epigenetic features of tumour genomes to finding a cure for cancers such as retinoblastoma.

## METHODS SUMMARY

Full details of sample acquisition, molecular and biochemical procedures, informatics and whole genome sequencing, and animal and drug studies are provided in Supplementary Information. The St Jude Children's Research Hospital Institutional Review Board approved experiments involving human subjects and informed consent was obtained from all subjects. For animal studies, all experiments were performed in accordance with federal guidelines and regulations. The St Jude Children's Research Hospital Institutional Animal Care and Use Committee approved all animal experiments. Lentiviral vectors (GIPZ with Lenti-SYK-9 #V3LHS-366147 and Lenti-SYK-6 #V3LHS-366143) encoding shRNAs against SYK were purchased from OpenBiosystems.

**Received 27 March; accepted 24 November 2011.**

**Published online 11 January 2012.**

1. Friend, S. H. *et al.* A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323,** 643–646 (1986).
2. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
3. Knudson, A. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68,** 820–823 (1971).
4. Manning, A. L., Longworth, M. S. & Dyson, N. J. Loss of pRB causes centromere dysfunction and chromosomal instability. *Genes Dev.* **24,** 1364–1376 (2010).
5. Hernando, E. *et al.* Rb inactivation promotes genomic instability by uncoupling cell cycle progression from mitotic control. *Nature* **430,** 797–802 (2004).
6. Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications – miswritten, misinterpreted and mis-erased in human cancers. *Nature Rev. Cancer* **10,** 457–469 (2010).
7. Lu, J., Ruhf, M. L., Perrimon, N. & Leder, P. A genome-wide RNA interference screen identifies putative chromatin regulators essential for E2F repression. *Proc. Natl Acad. Sci. USA* **104,** 9381–9386 (2007).
8. Benetti, R. *et al.* A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nature Struct. Mol. Biol.* **15,** 998 (2008).
9. Wen, H., Andrejka, L., Ashton, J., Karess, R. & Lipsick, J. S. Epigenetic regulation of gene expression by *Drosophila* Myb and E2F2-RBF via the Myb-MuvB/dREAM complex. *Genes Dev.* **22,** 601–614 (2008).
10. Bourgo, R. J. *et al.* SWI/SNF deficiency results in aberrant chromatin organization, mitotic failure, and diminished proliferative capacity. *Mol. Biol. Cell* **20,** 3192–3199 (2009).
11. Ley, T. J. *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456,** 66–72 (2008).
12. Corson, T. W. & Gallie, B. L. One hit, two hits, three hits, more? Genomic changes in the development of retinoblastoma. *Genes Chromosom. Cancer* **46,** 617–634 (2007).
13. Ganem, N. J., Godinho, S. A. & Pellman, D. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460,** 278–282 (2009).
14. Iovino, F., Lentini, L., Amato, A., &. Di Leonardo, A. RB acute loss induces centrosome amplification and aneuploidy in murine primary fibroblasts. *Mol. Cancer* **5,** 38 (2006).
15. Amato, A., Lentini, L., Schillaci, T., Iovino, F. & Di Leonardo, A. RNAi mediated acute depletion of retinoblastoma protein (pRb) promotes aneuploidy in human primary cells via micronuclei formation. *BMC Cell Biol.* **10,** 79 (2009).
16. McEvoy, J. *et al.* Coexpression of normally incompatible developmental pathways in retinoblastoma genesis. *Cancer Cell* **16,** 260–275 (2011).
17. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nature Rev. Cancer* **4,** 143–153 (2004).
18. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nature Genet.* **21,** 163–167 (1999).
19. Laird, P. W. Cancer epigenetics. *Hum. Mol. Genet.* **14** (suppl. 1), R65–R76 (2005).
20. Hansen, K. D. *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nature Genet.* **43,** 768–775 (2011).
21. Hahn, C. K. *et al.* Proteomic and genetic approaches identify Syk as an AML target. *Cancer Cell* **16,** 281–294 (2009).
22. Chen, L. *et al.* SYK-dependent tonic B-cell receptor signaling is a rational treatment target in diffuse large B-cell lymphoma. *Blood* **111,** 2230–2237 (2008).
23. Feldman, A. L. *et al.* Overexpression of Syk tyrosine kinase in peripheral T-cell lymphomas. *Leukemia* **22,** 1139–1143 (2008).
24. Young, R. M. *et al.* Mouse models of non-Hodgkin lymphoma reveal Syk as an important therapeutic target. *Blood* **113,** 2508–2516 (2009).
25. Weinblatt, M. E. *et al.* An oral spleen tyrosine kinase (Syk) inhibitor for rheumatoid arthritis. *N. Engl. J. Med.* **363,** 1303–1312 (2010).
26. Baudot, A. D. *et al.* The tyrosine kinase Syk regulates the survival of chronic lymphocytic leukemia B cells through PKCδ and proteasome-dependent regulation of Mcl-1 expression. *Oncogene* **28,** 3261–3273 (2009).
27. Suljagic, M. *et al.* The Syk inhibitor fostamatinib disodium (R788) inhibits tumor growth in the Eμ-TCL1 transgenic mouse model of CLL by blocking antigen-dependent B-cell receptor signaling. *Blood* **116,** 4894–4905 (2010).
28. Buchner, M. *et al.* Spleen tyrosine kinase inhibition prevents chemokine- and integrin-mediated stromal protective effects in chronic lymphocytic leukemia. *Blood* **115,** 4497–4506 (2010).
29. Gobessi, S. *et al.* Inhibition of constitutive and BCR-induced Syk activation downregulates Mcl-1 and induces apoptosis in chronic lymphocytic leukemia B cells. *Leukemia* **23,** 686–697 (2009).
30. Brennan, R. C. *et al.* Targeting the p53 pathway in retinoblastoma with subconjunctival Nutlin-3a. *Cancer Res.* **71,** 4205–4213 (2011).
31. Yan, Y. *et al.* Engraftment and growth of patient-derived retinoblastoma tumour in severe combined immunodeficiency mice. *Eur. J. Cancer* **36,** 221–228 (2000).
32. Squire, J., Gallie, B. L. & Phillips, R. A. A detailed analysis of chromosomal changes in heritable and non-heritable retinoblastoma. *Hum. Genet.* **70,** 291–301 (1985).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** The WGS was supported as part of the St Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project. We thank J. Lahti, J. Dalton and M. Valentine for help with FISH analysis, L. Holmfeldt, J. Zhang and M. Barbato for help with sample preparation, and V. Valentine for spectral karyotype analysis. We thank I. Qaddoumi, C. Rodriguez-Galindo and B. Haik for the support of the St Jude Children's Research Hospital retinoblastoma clinical research and W. Lei, D. McGoldrick, D. Alford, S. Espy, J. Obenauer and K. Johnson for assistance with data acquisition, handling and analysis. We thank A. McArthur and C. Guess for editing the manuscript, J. Temirov for help with sister chromatid analysis, J. Thurman for help with histology, F. Krafcik for help with cell culture, J. Wu and C. Billups for statistical analysis, J. Jeon for help with lentiviral preparations and C. Bradley for assistance with preclinical testing. This work was supported, in part, by Cancer Center Support (CA21765) from the NCI; grants to M.A.D. from the NIH (EY014867 and EY018599), the American Cancer Society and the Research to Prevent Blindness Foundation; and the American Lebanese Syrian Associated Charities. M.A.D. is a Howard Hughes Medical Institute Early Career Scientist. This work was also supported by an American Cancer Society Fellowship to A.L.M., the MGH Cancer Center Saltonstall Foundation Scholarship to N.J.D. and funding from AstraZeneca and NIH grants GM81607 and CA64402 to N.J.D.

**Author Contributions** M.A.D., C.A.B., J. McEvoy, J.F.-O., R.B. and A.L.M. designed the experiments. L.D. and J.Z. led the data analysis. M.A.D., L.D., J.Z., X.C., A.U., G.W., J.W., M.R., J. Ma, S.P., S.M., P.G., G.N., D.Z., C.L., R.S.F., L.L.F., X.H., D.J.D. and K.O. performed the data analysis and provided bioinformatics support. M.A.D., J.Z., A.U., C.A.B., J. McEvoy, J.F.-O., R.B. and A.L.M. prepared the figures and tables. C.A.B, J.M., J.F.-O., R.B. and A.L.M. performed the laboratory experiments. J.F.-O. created the xenograft. M.W. provided samples. J.E., S.S. and C.M. provided assistance with sample processing and handling. A.B. and D.E. provided pathology support. M.A.D., C.N., E.R.M., R.K.W. and J.R.D. supervised the project. M.A.D. wrote the manuscript. J.Z., C.A.B., J. McEvoy, J.F.-O., L.D., M.W., E.R.M., S.S., C.M. and J.R.D. read and commented on the manuscript. J.Z., L.D., X.C., A.U., G.W., J.W., M.R., J. Ma, J.E., S.S., C.M., S.P., S.M., P.G., G.N., D.Z., C.L., R.S.F., L.L.F., X.H., D.J.D., K.O., C.N., E.R.M., A.B., D.E., R.K.W., J.R.D. and M.A.D. are part of the St Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project.

**Author Information** The whole-genome sequence data and SNP 6.0 data have been deposited in the NCBI dbGaP under accession number phs000352.v1.p1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.A.D. (michael.dyer@stjude.org), J.R.D. (james.downing@stjude.org) or R.K.W. (rwilson@wustl.edu).

doi:10.1038/nature10728

# Structure of HDAC3 bound to co-repressor and inositol tetraphosphate

Peter J. Watson[1], Louise Fairall[1], Guilherme M. Santos[1] & John W. R. Schwabe[1]

**Histone deacetylase enzymes (HDACs) are emerging cancer drug targets. They regulate gene expression by removing acetyl groups from lysine residues in histone tails, resulting in chromatin condensation. The enzymatic activity of most class I HDACs requires recruitment into multi-subunit co-repressor complexes, which are in turn recruited to chromatin by repressive transcription factors. Here we report the structure of a complex between an HDAC and a co-repressor, namely, human HDAC3 with the deacetylase activation domain (DAD) from the human SMRT co-repressor (also known as NCOR2). The structure reveals two remarkable features. First, the SMRT-DAD undergoes a large structural rearrangement on forming the complex. Second, there is an essential inositol tetraphosphate molecule—D-myo-inositol-(1,4,5,6)-tetrakisphosphate (Ins(1,4,5,6)$P_4$)—acting as an 'intermolecular glue' between the two proteins. Assembly of the complex is clearly dependent on the Ins(1,4,5,6)$P_4$, which may act as a regulator—potentially explaining why inositol phosphates and their kinases have been found to act as transcriptional regulators. This mechanism for the activation of HDAC3 appears to be conserved in class I HDACs from yeast to humans, and opens the way to novel therapeutic opportunities.**

The acetylation of lysine residues in the tails of histone proteins plays an important role in the regulation of gene expression in eukaryotic cells[1,2]. The level of lysine acetylation is controlled through the opposing actions of histone acetyl transferases (HATs) and HDACs. Although chromatin is the best understood substrate for these enzymes, lysine acetylation is emerging as a general regulatory mechanism in a diverse array of cellular processes[3].

There are four classes of HDACs in mammalian cells (reviewed in ref. 4). Class I HDACs are zinc-dependent enzymes, and include HDAC1, 2, 3 and 8. Of these, only HDAC8 is a fully functional enzyme in isolation[5,6]. HDAC1, 2 and 3 require recruitment into large multi-subunit co-repressor complexes for maximal activity[7–12]. These co-repressor complexes bring about the repression of gene expression when recruited to repressive transcription factors, but also contribute to the 'resetting' of chromatin after rounds of transcriptional activation[13–16].

In recent years, HDACs have become important targets for the treatment of a number of cancers[17]. Cancer cell lines treated with HDAC inhibitors undergo terminal differentiation, growth arrest and/or apoptosis. Several HDAC inhibitors are at various stages in clinical trials, and two drugs, vorinostat and romidepsin, have been approved for the treatment of cutaneous T-cell lymphomas[18].

HDAC1 and 2 are found in three co-repressor complexes: NuRD[7,19], CoREST[20,21] and Sin3A[22,23]. In contrast, HDAC3 appears to be uniquely recruited to the SMRT complex (or to the homologous NCoR complex) where it interacts with a conserved DAD within SMRT or NCoR[8,10,12,24–26]. The DAD both recruits and activates HDAC3[10–12]. Recruitment of HDAC3 to the DAD is essential for repression by certain nuclear receptors and for the maintenance of normal circadian physiology[27–29]. It has been proposed that the assembly of the HDAC3 and SMRT-DAD requires a chaperone complex, because when these proteins are expressed in bacteria they do not form a complex[30].

The DAD contains an extended SANT-like domain with an amino-terminal DAD-specific motif. Deletion of this motif results in both loss of binding and failure to activate HDAC3[11].We have previously reported the structure of the isolated DAD from SMRT[31]. This revealed that part of the DAD-specific motif forms an extra helix that is folded against the three helices of the SANT domain to form a four-helix bundle. The N-terminal portion of the DAD-specific motif is unstructured in solution[31].

Here we report the structure of HDAC3 in complex with SMRT-DAD. This structure not only reveals the specificity and mechanism through which SMRT-DAD recruits and activates HDAC3, but also identifies Ins(1,4,5,6)$P_4$ as a key component of the complex that has the potential to regulate assembly of HDACs with their co-repressors.

## Overall architecture of the complex

As HDAC3 and SMRT-DAD do not form a complex when expressed in bacterial cells, full-length HDAC3 and Flag-tagged SMRT-DAD (amino acids 389–480) were expressed in suspension-grown mammalian HEK293 cells. The complex remained tightly associated during a three-step purification, including size exclusion chromatography (Supplementary Fig. 1). Interestingly, at salt concentrations higher than 50 mM, the complex dissociated and the HDAC3 predominately took the form of aggregate or oligomers (Supplementary Fig. 2). Such oligomerization of HDAC3 has also been reported by others[32]. During crystallization, the HDAC3 tail is proteolysed but HDAC3 remains bound to SMRT-DAD and retains deacetylase activity (Supplementary Fig. 3). Small crystals (15 μm; Supplementary Fig. 1) diffracted to 2 Å and the structure was solved by molecular replacement with an HDAC8 structure[33]. Model building and refinement yielded an excellent map to 2.1 Å resolution with clear density for both HDAC3 and SMRT-DAD (Supplementary Fig. 1).

Overall, the HDAC3 structure is similar to the previously determined class I HDAC structures of HDAC8[34] and HDAC2[35], and consists of an eight-stranded parallel β-sheet surrounded by a number of α-helices. The active site lies at the base of a tunnel leading from the surface of the protein. A solvent-exposed tyrosine residue is located on the surface of the enzyme immediately adjacent to the active site tunnel. This tyrosine is unique to HDAC3, and it seems that this residue might interact with the substrate and possibly contribute to substrate specificity (Supplementary Fig. 4).

---

[1]Henry Wellcome Laboratories of Structural Biology, Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK.

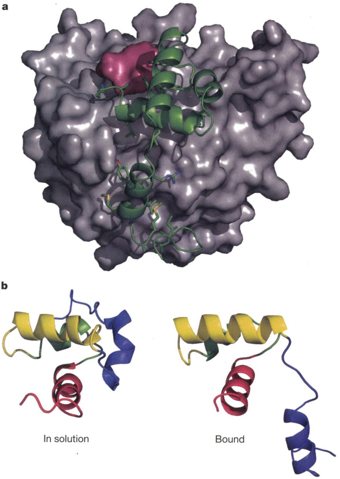
## Structural rearrangement of SMRT–DAD

On forming a complex with HDAC3, the N-terminal helix of the DAD undergoes a major structural rearrangement (from that seen in the NMR structure), such that it no longer forms part of the core structure, but lies along the surface of HDAC3, making extensive intermolecular interactions (Fig. 1a, b). Along with a further extended region, this DAD-specific motif buries a surface of 1,178 $\text{Å}^2$. The remaining three-helix bundle resembles a canonical SANT domain, and buries a further 1,160 $\text{Å}^2$ at the interface with HDAC3.

The SMRT-DAD interacts with the N-terminal region of HDAC3 (residues 9–49) that forms helix H1, loop L2, helix H2 and strand S2. This region differs between HDAC8 and HDAC3. Indeed, in HDAC3, helix H1 is distorted and can only be considered to be a pseudo-helix. These differences may explain why HDAC8 is active in the absence of an interacting co-repressor.

## An essential role for Ins(1,4,5,6)$P_4$

At the earliest stages of refinement, the electron density difference map revealed a well-ordered small molecule bound at the interface between HDAC3 and the DAD (Fig. 2a). The electron density was sufficiently well defined that the small molecule could be readily identified as inositol tetraphosphate. During further refinement, it could be unambiguously assigned as D-myo-inositol-1,4,5,6-tetrakisphosphate (based on the axial orientation of the hydroxyl group on carbon 2), referred to here as Ins(1,4,5,6)$P_4$ (Fig. 2b).



In solution

Bound

**Figure 1 | Structure of the HDAC3–SMRT-DAD complex.** **a**, Interaction of SMRT-DAD (green ribbon) with HDAC3 (grey surface). The Ins(1,4,5,6)$P_4$ at the interface is shown as a raspberry coloured surface. Side chains in the DAD that mediate interaction with HDAC3 and Ins(1,4,5,6)$P_4$ are shown as sticks. **b**, Structure of the DAD domain in solution (left; PDB code 1XC5) compared with that bound to HDAC3 (right); helices are individually coloured to facilitate comparison.

The Ins(1,4,5,6)$P_4$ molecule is sandwiched between HDAC3 and the DAD, in a highly basic pocket, making extensive contacts to both proteins, burying a surface of 407 $\text{Å}^2$ (Fig. 2c, d). HDAC3 and SMRT-DAD each contribute five hydrogen bonds and salt bridges to the Ins(1,4,5,6)$P_4$ (His 17, Gly 21, Lys 25, Arg 265, Arg 301 and Lys 449, Tyr 470, Tyr 471, Lys 474 and Lys 475, respectively) (Fig. 2e).

It is notable that the Ins(1,4,5,6)$P_4$ is sufficiently tightly bound to the complex that it is retained through the entire purification process. It is also striking that the binding is highly specific for Ins(1,4,5,6)$P_4$, as the electron density shows that the ligand is uniquely Ins(1,4,5,6)$P_4$ rather than a mixture of inositol phosphates.

A careful examination of the structure suggests that Ins(1,4,5,6)$P_4$ binding is an essential requirement for the interaction between SMRT and HDAC3, acting as a 'intermolecular glue' that cements the complex together. If the Ins(1,4,5,6)$P_4$ ligand were not present, then the many basic residues on either side of the binding cleft would probably prevent interaction through charge repulsion. Indeed, at the base of the cleft, the ζ-nitrogen atoms of SMRT Lys 449 and of HDAC3 Lys 25 are just 4.4 Å from each other. Consequently, the assembly of the three-way SMRT–HDAC3–Ins(1,4,5,6)$P_4$ complex appears to be mutually interdependent, such that both SMRT-DAD and Ins(1,4,5,6)$P_4$ are required for activation of the HDAC3 enzymatic functionality.

These conclusions are supported by our previous mutagenesis study of SMRT-DAD, in which we looked at the effect of mutations on both the interaction with, and the deacetylase activity of, HDAC3. All mutations that failed to activate HDAC3 also abolished, or significantly impaired, interaction. These include mutations of Lys 449, Tyr 470 and Tyr 471, which play a key role in binding Ins(1,4,5,6)$P_4$ (ref. 31). The requirement for Ins(1,4,5,6)$P_4$ to enable complex formation may contribute to the explanation as to why recombinant HDAC3 and SMRT-DAD expressed in bacteria fail to interact, as bacteria probably do not contain sufficient Ins(1,4,5,6)$P_4$ to support complex formation.

Having discovered that Ins(1,4,5,6)$P_4$ plays a key role in HDAC3 activation, we asked whether inositol phosphates might contribute to the assembly and activation of other class I HDAC complexes. Significantly, the residues that mediate interaction with Ins(1,4,5,6)$P_4$ and the SANT domain from SMRT are conserved in both HDAC1 and 2, but not in HDAC8 (Fig. 3a), fitting with the observation that only HDAC8 is fully active in isolation.

Similarly, the co-repressor partners for HDAC1 and 2 (that is, MTA1, 2 and 3 and CoREST1, 2 and 3) contain SANT domains that are analogous to SMRT-DAD. The key Ins(1,4,5,6)$P_4$ binding residues are almost entirely conserved (Fig. 3b). It is likely that the specificity for the particular HDAC is conferred by the region N-terminal to the SANT domain: that is, the DAD specific motif in SMRT/NCoR and the ELM2 domains in the MTA and CoREST proteins.

Taken together, this suggests that inositol phosphate activation of class I HDACs is a general mechanism and, given that the key residues are also found in the yeast HDAC, Rpd3, and the Snt1 co-repressor, it would appear that this mechanism is also evolutionarily conserved.

## Mechanism of activation of HDAC3

The HDAC3–SMRT-DAD structure provides insight into the mechanism through which the DAD and Ins(1,4,5,6)$P_4$ contribute to the activation of HDAC3. In the crystal, the active site of HDAC3 resembles a product complex (Fig. 4a, b). An acetate molecule (present during purification) is bound at the active site, making hydrogen bonds to the catalytic zinc and side chains of Ty 298 and His 134. Furthermore, a methionine side chain (from an adjacent SMRT-DAD in the crystal lattice) is bound in the active site tunnel, mimicking a lysine residue.

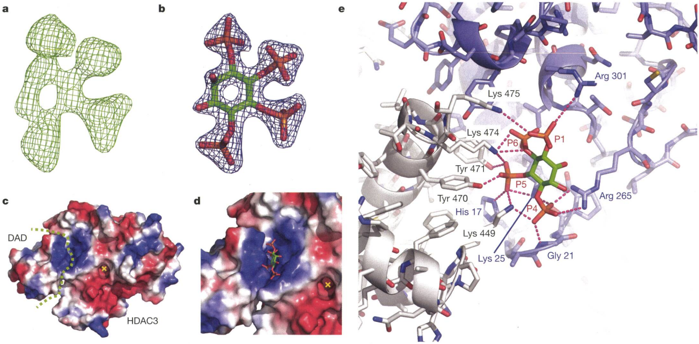
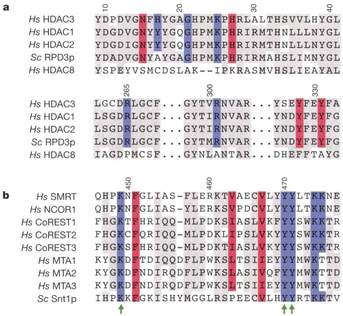


**Figure 2 | Ins(1,4,5,6)$P_4$ binding to the HDAC3–SMRT-DAD complex.** **a**, A striking feature in the difference electron density map ($F_o - F_c$ at $3\sigma$, green mesh) observed following molecular replacement. **b**, Electron density corresponding to the Ins(1,4,5,6)$P_4$ ligand following refinement ($2F_o - F_c$ at $2.25\sigma$, blue mesh). The Ins(1,4,5,6)$P_4$ is shown as a stick model. **c**, Electrostatic surface representation of the HDAC3–SMRT-DAD complex (red indicates negative potential and blue positive). A strikingly basic pocket is located at the HDAC3–SMRT-DAD interface (indicated by a dashed green line). The active site pocket of HDAC3 is indicated by a yellow cross. The Ins(1,4,5,6)$P_4$ is not shown for clarity. **d**, Ins(1,4,5,6)$P_4$ (shown as sticks) binding in the basic pocket at the HDAC3–SMRT-DAD interface. **e**, Detailed interactions of Ins(1,4,5,6)$P_4$ (sticks) with HDAC3 (blue) and SMRT-DAD (grey). Side chains from SMRT and HDAC3 which make contacts with the Ins(1,4,5,6)$P_4$ are labelled. The phosphate groups of the Ins(1,4,5,6)$P_4$ are labelled P1–6. Hydrogen bonds and/or salt bridges are indicated by dashed lines.


a
10
20
30
40
265
300
330
b
450
460
470

**a**

| | |
|---|---|
| *Hs* HDAC3 | YDPDVGNFHYGAGHPMKPHRLALTHSVVLHYGL |
| *Hs* HDAC1 | YDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGL |
| *Hs* HDAC2 | YDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGL |
| *Sc* RPD3p | YDADVGNYAYGAGHPMKPHRIRMAHSLIMNYGL |
| *Hs* HDAC8 | YSPEYVSMCDSLAK--IPKRASMVHSLIEAYAL |

| | |
|---|---|
| *Hs* HDAC3 | LGCDRLGCF...GYTVRNVAR...YSEYFEYFA |
| *Hs* HDAC1 | LSGDRLGCF...GYTIRNVAR...YNDYFEYFG |
| *Hs* HDAC2 | LSGDRLGCF...GYTIRNVAR...YNDYFEYFG |
| *Sc* RPD3p | LSGDRLGCF...GYTMRNVAR...YNDYFEYFG |
| *Hs* HDAC8 | IAGDPMCSF...GYNLANTAR...DHEFFTAYG |

**b**

| | |
|---|---|
| *Hs* SMRT | QHPKNFGLIAS-FLERKTVAECVLYYYLTKKNE |
| *Hs* NCOR1 | QHPKNFGLIAS-YLERKSVPDCVLYYYLTKKNE |
| *Hs* CoREST1 | FHGKTFHRIQQ-MLPDKSIASLVKFYYSWKKTR |
| *Hs* CoREST2 | FHGKCFQRIQQ-MLPDKLIPSLVKYYYSWKKTR |
| *Hs* CoREST3 | FHGKSFHRIQQ-MLPDKTIASLVKYYYSWKKTR |
| *Hs* MTA1 | KYGKDFTDIQQDFLPWKSLTSIIEYYYMWKTTD |
| *Hs* MTA2 | KYGKDFNDIRQDFLPWKSLASIVQFYYMWKTTD |
| *Hs* MTA3 | KYGKDFNDIRQDFLPWKSLTSIIEYYYMWKTTD |
| *Sc* Snt1p | IHPKKFGKISHYMGGLRSPEECVLHYYRTKKTV |

**Figure 3 | Sequence conservation suggests that class I HDACs, from yeast to man, require inositol phosphates for assembly and activation.** **a**, Alignments of key regions of class I HDACs from *Homo sapiens* (*Hs*) and *Saccharomyces cerevisiae* (*Sc*). Residues that mediate interaction with Ins(1,4,5,6)$P_4$ and SMRT-DAD are highlighted in blue and red, respectively. Other conserved residues are highlighted in grey. **b**, Alignment of SANT domains from known interaction partners for class I HDACs. Residues that mediate interaction with Ins(1,4,5,6)$P_4$ and HDAC3 are highlighted in blue and red, respectively. Green arrows indicate residues that impair HDAC3 recruitment and activation when mutated to alanine[31].

The binding surfaces for DAD and Ins(1,4,5,6)$P_4$ are located to one side of the HDAC3 active site, involving pseudo helix H1 and loops L1 and L6 (Fig. 4a–c). We propose that changes in both conformation and dynamics occur when DAD and Ins(1,4,5,6)$P_4$ bind to HDAC3, and that these facilitate substrate access to the active site, resulting in enhanced enzyme activity.

There is a key interaction between Ins(1,4,5,6)$P_4$ and Arg 265 in loop L6 (coloured orange in Fig. 4b). This loop seems to be very important for access to the active site, because Leu 266 forms one wall of the active site tunnel, and in the absence of the Ins(1,4,5,6)$P_4$ this loop is likely to be relatively mobile.

Comparison of the HDAC3 structure with that of HDAC8, which does not require activation by complex formation, reveals that HDAC8 differs significantly in the region where HDAC3 interacts with SMRT-DAD and Ins(1,4,5,6)$P_4$ (Fig. 4d). In HDAC8, the equivalent helix to pseudo helix H1 in HDAC3 has a regular helical structure, loop L1 is two amino acids shorter and loop L6 contains a proline residue that partly orientates the loop away from the active site (Fig. 4d). We suggest that together these differences give substrate better access to that active site of HDAC8 than would be possible in the uncomplexed HDAC3. The pattern of crystallographic temperature factors for the various structures supports this interpretation (Supplementary Fig. 5).

To test the importance of Arg 265 as well as loops L1 and L6, we co-expressed mutant HDAC3 constructs with SMRT-DAD in mammalian cells. Strikingly, all the designed mutations resulted not only in total loss of deacetylase activity (Fig. 5a) but also abolished interaction with SMRT-DAD (Fig. 5b). These findings support our interpretation of the importance of these residues in the activation of HDAC3.

To support the conclusions that Ins(1,4,5,6)$P_4$ is essential for HDAC3 activity and to test the specificity for Ins(1,4,5,6)$P_4$, we sought to establish an *in vitro* reconstitution assay. We expressed

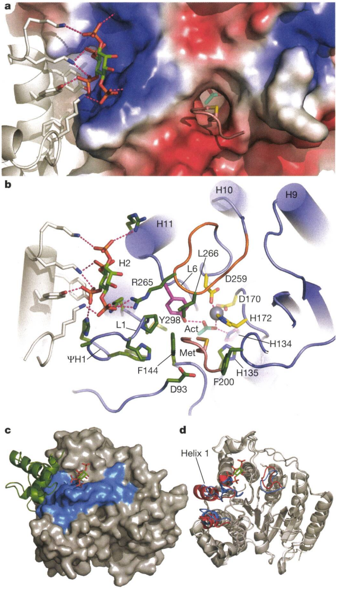


**Figure 4 | Mechanism of activation of HDAC3 by binding SMRT-DAD and Ins(1,4,5,6)$P_4$.** **a**, SMRT-DAD (grey cartoon) and Ins(1,4,5,6)$P_4$ bind adjacent to the HDAC3 (charged surface representation) active site. Acetate and a methionine (lysine mimic) are located in the active site. **b**, Details of the HDAC3 active site. Key residues and loops are labelled—see text for details (hydrogen bonds and/or salt bridges are indicated by dashed lines). ΨH1, pseudo helix H1. **c**, Pseudo helix H1 and loops L1 and L6 are shown in blue on the surface of HDAC3 (grey). These regions are influenced or stabilized by binding of SMRT-DAD (green) and Ins(1,4,5,6)$P_4$ (shown as sticks). **d**, Comparison of the structures of HDAC3 and HDAC8. Regions of significant difference are coloured blue (HDAC3) and red (HDAC8). Helix 1 is indicated.

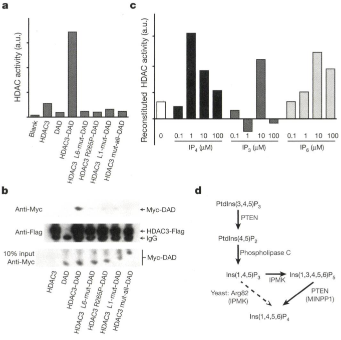


**Figure 5 | Exploring the role of Ins(1,4,5,6)$P_4$ in complex assembly and HDAC3 activation.** **a**, Effect of HDAC3 mutations on deacetylase activity of complexes with SMRT-DAD. Residues are mutated to their equivalents in HDAC8. Loop 1 mutation (L1-mut) is H17C/G21A/K25I. Loop 6 mutation (L6-mut) is R264P/L265M. 'Mut-all' is H17C/G21A/K25I/R264P/L265M/R301A. **b**, Immunoblots showing that HDAC3 mutations perturb interaction with SMRT-DAD. Recombinant proteins (Myc-DAD, HDAC3-Flag) and the IgG heavy chain are indicated. **c**, Deacetylase assays of HDAC3-Flag reconstituted with bacterially expressed SMRT-DAD in the presence of various inositol phosphates. IP4, IP3 and IP6 are Ins(1,4,5,6)$P_4$, Ins(1,4,5)$P_3$ and Ins(1,2,3,4,5,6)$P_6$, respectively. **d**, Synthesis pathway for Ins(1,4,5,6)$P_4$ from phosphatidylinositol-(3,4,5)-trisphosphate (PtdIns(3,4,5)$P_3$) and phosphatidylinositol-(4,5)-bisphosphate (PtdIns(4,5)$P_2$). Yeast Arg 82 converts Ins(1,4,5)$P_3$ to Ins(1,4,5,6)$P_4$. In mammals, both IPMK and PTEN are required to make Ins(1,4,5,6)$P_4$.

and purified HDAC3, Flag-tagged at the carboxy terminus, from mammalian cells and incubated the anti-Flag resin with bacterially expressed SMRT-DAD and various concentrations of Ins(1,4,5,6)$P_4$, Ins(1,4,5)$P_3$ and Ins(1,2,3,4,5,6)$P_6$ (Fig. 5c). HDAC activity was very sensitive to inositol phosphate concentration and higher concentrations do not support reconstitution. This can be explained through high levels of free inositol phosphate competing with complex formation. This is analogous to moderate monovalent salt concentrations causing complex dissociation during purification (Supplementary Fig. 2).

A similar pattern of activation was seen using both Ins(1,4,5)$P_3$ and Ins(1,2,3,4,5,6)$P_6$, but in both cases a tenfold higher concentration was required to achieve a similar level of HDAC activity to that seen with Ins(1,4,5,6)$P_4$. This supports the interpretation that Ins(1,4,5,6)$P_4$ is the physiologically relevant assembly partner.

## Does Ins(1,4,5,6)$P_4$ regulate HDACs?

The finding that Ins(1,4,5,6)$P_4$ is essential for the assembly of class 1 HDAC repression complexes raises the question as to whether it is a signalling molecule with a role in regulating complex assembly, or if it is simply an essential structural cofactor.

The fact that Ins(1,4,5,6)$P_4$ remained bound during purification would seem to suggest that it is a structural cofactor. However, to retain an intact complex it was necessary to use low ionic strength buffers. At physiological ionic strengths, the Ins(1,4,5,6)$P_4$ is likely to dissociate readily and thus could regulate complex assembly.

Importantly, several studies have previously implicated Ins(1,4,5,6)$P_4$ as a regulator of gene expression and chromatin remodelling—although no molecular mechanism was identified. Yeast Arg 82 is a transcriptional regulator of genes involved in arginine metabolism. It has been shown[36] that Arg 82 is an inositol phosphate kinase that converts Ins(1,4,5)$P_3$ to Ins(1,4,5,6)$P_4$ and that kinase activity is required for its role in transcriptional regulation (Fig. 5d). Several further studies have supported the importance of Arg 82 kinase activity in transcriptional regulation and chromatin remodelling[37–39], together suggesting that Ins(1,4,5,6)$P_4$ has an important regulatory role in yeast.

IPMK, the mammalian homologue of Arg 82, has been reported to phosphorylate Ins(1,4,5)$P_3$ to form Ins(1,3,4,5,6)$P_4$ and then

Ins(1,3,4,5,6)$P_5$ (refs 40, 41). Consequently, in mammalian cells, a phosphatase is required to generate Ins(1,4,5,6)$P_4$. The most likely enzyme is the well-known tumour suppressor gene PTEN[42,43], which is known to be active in the nucleus and to play a role in chromosome stability (reviewed in ref. 44). It is tempting to speculate that loss of HDAC complex function might be one of the routes through which PTEN mutations contribute to oncogenesis.

A final important question is whether the levels of Ins(1,4,5,6)$P_4$ are regulated in the cell. This is not easy to establish, as regulation could be temporally and/or spatially compartmentalized. However there is some evidence that both PTEN and $InsP_4$ levels change with progression through the cell cycle[45]. It is also possible that changing levels of Ins(1,4,5,6)$P_4$ may contribute to the circadian regulation of HDAC3 activity[46]. Further research will be needed to clarify these important questions.

## Discussion

We present the structure of a histone deacetylase in complex with its activating co-repressor. The structure reveals a striking structural rearrangement of the co-repressor and an unexpected inositol tetraphosphate molecule (Ins(1,4,5,6)$P_4$) acting as an 'intermolecular glue', contributing to the stabilization and activation of HDAC3.

Sequence conservation suggests that Ins(1,4,5,6)$P_4$ plays a key role in co-repressor assembly and activation of class I HDACs from yeast to humans and is likely to act as a regulator of HDAC complex assembly. Although it remains to be established how the availability of Ins(1,4,5,6)$P_4$ might be controlled, the requirement of Ins(1,4,5,6)$P_4$ for co-repressor HDAC assembly presents novel opportunities for therapeutic intervention that may complement existing HDAC inhibitors. It may be possible to develop molecules that target the Ins(1,4,5,6)$P_4$ binding site itself, but it may also be possible to target the enzymes responsible for Ins(1,4,5,6)$P_4$ synthesis.

## METHODS SUMMARY

HDAC3 and Flag-tagged SMRT-DAD (residues 389–480) were expressed using transient transfection in suspension grown HEK293F cells. The HDAC3–SMRT-DAD complex was purified by Flag affinity chromatography, followed by TEV protease cleavage and size exclusion chromatography. Crystals of the HDAC3–DAD complex were grown by sitting drop vapour diffusion with a 10% propan-2-ol precipitant. The structure was determined by molecular replacement based on HDAC8 (PDB code 3EW8). Ligands, including Ins(1,4,5,6)$P_4$, zinc, potassium, acetate and glycerol were added during the refinement process.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 5 July; accepted 23 November 2011.**
**Published online 9 January 2012.**

1. Pogo, B. G., Allfrey, V. G. & Mirsky, A. E. RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc. Natl Acad. Sci. USA* **55,** 805–812 (1966).
2. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12,** 599–606 (1998).
3. Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840 (2009).
4. Hildmann, C., Riester, D. & Schwienhorst, A. Histone deacetylases — an important class of cellular regulators with a variety of functions. *Appl. Microbiol. Biotechnol.* **75,** 487–497 (2007).
5. Hu, E. *et al.* Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. *J. Biol. Chem.* **275,** 15254–15264 (2000).
6. Lee, H., Rezai-Zadeh, N. & Seto, E. Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A. *Mol. Cell. Biol.* **24,** 765–773 (2004).
7. Zhang, Y. *et al.* Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13,** 1924–1935 (1999).
8. Li, J. *et al.* Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J.* **19,** 4342–4350 (2000).
9. Lechner, T. *et al.* Sds3 (suppressor of defective silencing 3) is an integral component of the yeast Sin3·Rpd3 histone deacetylase complex and is required for histone deacetylase activity. *J. Biol. Chem.* **275,** 40961–40966 (2000).
10. Zhang, J., Kalkum, M., Chait, B. T. & Roeder, R. G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol. Cell* **9,** 611–623 (2002).
11. Guenther, M. G., Barak, O. & Lazar, M. A. The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol. Cell. Biol.* **21,** 6091–6101 (2001).
12. Wen, Y. D. *et al.* The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proc. Natl Acad. Sci. USA* **97,** 7202–7207 (2000).
13. Métivier, R. *et al.* Estrogen receptor-α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115,** 751–763 (2003).
14. Hoberg, J. E., Yeung, F. & Mayo, M. W. SMRT derepression by the IκB kinase α: a prerequisite to NF-κB transcription and survival. *Mol. Cell* **16,** 245–255 (2004).
15. Billin, A. N., Thirlwell, H. & Ayer, D. E. β-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator. *Mol. Cell. Biol.* **20,** 6882–6890 (2000).
16. Guan, H.-P., Ishizuka, T., Chui, P. C., Lehrke, M. & Lazar, M. A. Corepressors selectively control the transcriptional activity of PPARγ in adipocytes. *Genes Dev.* **19,** 453–461 (2005).
17. Marks, P. A. & Breslow, R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nature Biotechnol.* **25,** 84–90 (2007).
18. Wagner, J. M., Hackanson, B., Lübbert, M. & Jung, M. Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin. Epigenet.* **1,** 117–136 (2010).
19. Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S. & Reinberg, D. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95,** 279–289 (1998).
20. Hakimi, M.-A. *et al.* A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc. Natl Acad. Sci. USA* **99,** 7420–7425 (2002).
21. Humphrey, G. W. *et al.* Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J. Biol. Chem.* **276,** 6817–6824 (2001).
22. Laherty, C. D. *et al.* Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* **89,** 349–356 (1997).
23. Heinzel, T. *et al.* A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature* **387,** 43–48 (1997).
24. Guenther, M. G. *et al.* A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes Dev.* **14,** 1048–1057 (2000).
25. Yoon, H.-G. *et al.* Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* **22,** 1336–1346 (2003).
26. Oberoi, J. *et al.* Structural basis for the assembly of the SMRT/NCoR core transcriptional repression machinery. *Nature Struct. Mol. Biol.* **18,** 177–184 (2011).
27. You, S.-H., Liao, X., Weiss, R. E. & Lazar, M. A. The interaction between nuclear receptor corepressor and histone deacetylase 3 regulates both positive and negative thyroid hormone action *in vivo*. *Mol. Endocrinol.* **24,** 1359–1367 (2010).
28. Ishizuka, T. & Lazar, M. A. The nuclear receptor corepressor deacetylase activating domain is essential for repression by thyroid hormone receptor. *Mol. Endocrinol.* **19,** 1443–1451 (2005).
29. Yin, L. *et al.* Rev-erbα, a heme sensor that coordinates metabolic and circadian pathways. *Science* **318,** 1786–1789 (2007).
30. Guenther, M. G. Assembly of the SMRT-histone deacetylase 3 repression complex requires the TCP-1 ring complex. *Genes Dev.* **16,** 3130–3135 (2002).
31. Codina, A. *et al.* Structural insights into the interaction and activation of histone deacetylase 3 by nuclear receptor corepressors. *Proc. Natl Acad. Sci. USA* **102,** 6009–6014 (2005).
32. Yang, W.-M., Tsai, S.-C., Wen, Y.-D., Fejer, G. & Seto, E. Functional domains of histone deacetylase-3. *J. Biol. Chem.* **277,** 9447–9454 (2002).
33. Dowling, D. P., Gantt, S. L., Gattis, S. G., Fierke, C. A. & Christianson, D. W. Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors. *Biochemistry* **47,** 13554–13563 (2008).
34. Somoza, J. R. *et al.* Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure* **12,** 1325–1334 (2004).
35. Bressi, J. C. *et al.* Exploration of the HDAC2 foot pocket: synthesis and SAR of substituted N-(2-aminophenyl)benzamides. *Bioorg. Med. Chem. Lett.* **20,** 3142–3145 (2010).
36. Odom, A. R. A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control. *Science* **287,** 2026–2029 (2000).
37. Steger, D. J. Regulation of chromatin remodeling by inositol polyphosphates. *Science* **299,** 114–116 (2003).
38. El Alami, M., Messenguy, F., Scherens, B. & Dubois, E. Arg82p is a bifunctional protein whose inositol polyphosphate kinase activity is essential for nitrogen and PHO gene expression but not for Mcm1p chaperoning in yeast. *Mol. Microbiol.* **49,** 457–468 (2003).
39. Shen, X., Xiao, H., Ranallo, R., Wu, W.-H. & Wu, C. Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* **299,** 112–114 (2003).
40. Saiardi, A. *et al.* Mammalian inositol polyphosphate multikinase synthesizes inositol 1,4,5-trisphosphate and an inositol pyrophosphate. *Proc. Natl Acad. Sci. USA* **98,** 2306–2311 (2001).

41. Nalaskowski, M. M., Deschermeier, C., Fanick, W. & Mayr, G. W. The human homologue of yeast ArgRIII protein is an inositol phosphate multikinase with predominantly nuclear localization. *Biochem. J.* **366,** 549–556 (2002).
42. Craxton, A., Caffrey, J. J., Burkhart, W., Safrany, S. T. & Shears, S. B. Molecular cloning and expression of a rat hepatic multiple inositol polyphosphate phosphatase. *Biochem. J.* **328,** 75–81 (1997).
43. Caffrey, J. J., Darden, T., Wenk, M. R. & Shears, S. B. Expanding coincident signaling by PTEN through its inositol 1,3,4,5,6-pentakisphosphate 3-phosphatase activity. *FEBS Lett.* **499,** 6–10 (2001).
44. Di Cristofano, A. & Pandolfi, P. P. The multiple roles of PTEN in tumor suppression. *Cell* **100,** 387–390 (2000).
45. Mattingly, R. R., Stephens, L. R., Irvine, R. F. & Garrison, J. C. Effects of transformation with the v-src oncogene on inositol phosphate metabolism in rat-1 fibroblasts. D-myo-inositol 1,4,5,6-tetrakisphosphate is increased in v-src-transformed rat-1 fibroblasts and can be synthesized from D-myo-inositol 1,3,4-trisphosphate in cytosolic extracts. *J. Biol. Chem.* **266,** 15144–15153 (1991).
46. Feng, D. *et al.* A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. *Science* **331,** 1315–1319 (2011).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank R. Owen and the other beamline staff at DIAMOND I24 for help with data collection; J. Goodchild for help with biochemical experiments; and S. Cowley, M. Lazar, P. Moody, L. Nagy, P. Tontonoz and S. Shears for literature and discussions. This work was supported by the Wellcome Trust (grant WT085408).

**Author Contributions** P.J.W. expressed, purified and crystallized the protein and performed the biochemical studies. P.J.W., L.F. and J.W.R.S. performed the structural determination and wrote the paper. G.M.S. and L.F. performed early expression/purification trials in insect cells. J.W.R.S. conceived the study.

**Author Information** Atomic coordinates and structure factors are deposited in the Protein Data Bank under accession number 4A69. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.W.R.S. (john.schwabe@le.ac.uk).

doi:10.1038/nature10669

# Gravitational detection of a low-mass dark satellite galaxy at cosmological distance

S. Vegetti[1], D. J. Lagattuta[2], J. P. McKean[3], M. W. Auger[4], C. D. Fassnacht[2] & L. V. E. Koopmans[5]

**The mass function of dwarf satellite galaxies that are observed around Local Group galaxies differs substantially from simulations[1–5] based on cold dark matter: the simulations predict many more dwarf galaxies than are seen. The Local Group, however, may be anomalous in this regard[6,7]. A massive dark satellite in an early-type lens galaxy at a redshift of 0.222 was recently found[8] using a method based on gravitational lensing[9,10], suggesting that the mass fraction contained in substructure could be higher than is predicted from simulations. The lack of very low-mass detections, however, prohibited any constraint on their mass function. Here we report the presence of a $(1.9 \pm 0.1) \times 10^8 M_{\odot}$ dark satellite galaxy in the Einstein ring system JVAS B1938+666 (ref. 11) at a redshift of 0.881, where $M_{\odot}$ denotes the solar mass. This satellite galaxy has a mass similar to that of the Sagittarius[12] galaxy, which is a satellite of the Milky Way. We determine the logarithmic slope of the mass function for substructure beyond the local Universe to be $1.1^{+0.6}_{-0.4}$, with an average mass fraction of $3.3^{+3.6}_{-1.8}$ per cent, by combining data on both of these recently discovered galaxies. Our results are consistent with the predictions from cold dark matter simulations[13–15] at the 95 per cent confidence level, and therefore agree with the view that galaxies formed hierarchically in a Universe composed of cold dark matter.**

The gravitational lens system JVAS B1938+666 (ref. 11) has a bright infrared background galaxy at redshift $z = 2.059$ (ref. 16), which is gravitationally lensed into an almost complete Einstein ring of diameter ~0.9 arcsec by a massive elliptical galaxy at redshift $z = 0.881$ (ref. 17). This bright, highly magnified Einstein ring makes this system an excellent candidate in which to search for surface brightness anomalies caused by very low-mass (dark matter) substructure in the halo around the high-redshift elliptical lens galaxy. The presence of a low-mass substructure (for example a luminous or dark satellite galaxy) in the lens galaxy can introduce a localized perturbation of the arc surface brightness distribution. Owing to the multiplicity of the gravitationally lensed images that form these arcs, these surface brightness 'anomalies' can be analysed using a pixelated-lens modelling technique and used to gravitationally detect and quantify the total mass and position of the substructure, down to masses as low as ~0.1% of the mass of the lens inside the Einstein radius[9,10]. The lens system was imaged at 1.6 and 2.2 μm using the Near Infrared Camera 2 on the W. M. Keck 10-m telescope in June 2010. The adaptive optics system was used to correct the incoming wavefront for the blurring induced by the atmosphere, providing a nearly diffraction-limited point spread function with a full-width at half-maximum of ~70 mas. Further details of the data sets and their image processing can be found in Supplementary Information.

A smooth parametric model for the lens potential was constrained by using the surface brightness emission from the Einstein ring of each data set independently, having first removed the smooth light contribution from the lensing galaxy. We represented the mass model using an ellipsoidal power law, $\rho(r) \propto r^{-\gamma}$, where $\rho(r)$ is the combined luminous and dark matter density as a function of the ellipsoidal radius, $r$. The best-fitting model was then fixed and further refined using local potential corrections defined on a regular grid, which are translated into surface density corrections using the Laplace operator. We found for both the 1.6- and the 2.2-μm adaptive optics data sets that there was a significant positive density correction, which indicated the presence of a mass substructure (Fig. 1 and Supplementary Information). Directly from the pixelated potential correction, we measured a substructure mass of $\sim 1.7 \times 10^8 M_{\odot}$ inside a projected radius of 600 pc around the density peak.

As an independent test, we repeated the analysis of the 2.2-μm data set, which had the highest-significance positive density correction, with different models of the point spread function, different data reduction techniques, different rotations of the lensed images, different models for the lens galaxy surface brightness subtraction and different resolutions for the reconstructed source. We also analysed an independent data set taken at 1.6 μm with the Near Infrared Camera and Multi-Object Spectrograph on board NASA's Hubble Space Telescope. In total, we tested fourteen different models and three different data sets that all independently led to the detection of a positive density correction at the same spatial position, although with varying levels of significance (Supplementary Information). Differential extinction across the gravitational arc could also produce a surface brightness anomaly. However, the colour of the arc was found to be consistent around and at the location of substructure, ruling out the possibly that dust affected our results.

We used an analytic model to determine the mass and the statistical significance of the substructure in the context of a physical model[9,10]. In this analytic approach, a truncated pseudo-Jaffe model was used to parameterize the substructure mass and position, giving three extra free parameters. To obtain a good fit to the observed surface brightness distribution of the lensed background source at 2.2 μm, we found that a substructure was required at a position consistent with the positive density correction detected above. Assuming the substructure to be situated in the plane of the lens galaxy, we objectively compared the smooth and the substructure parametric models in terms of the Bayesian evidence, which is the probability of the data given the model (marginalized over all parameters). Computing the marginalized evidence involved integrating over the multidimensional parameter space within predefined priors. In particular, we assumed that the substructure was equally likely to be located at any point in the lens plane and to have a mass between $4.0 \times 10^6 M_{\odot}$ and $4.0 \times 10^9 M_{\odot}$, the mass range in which comparison with simulations is possible[13,14]. See Supplementary Information for further details on the Bayesian evidence. We found that our nominal model, with a substructure of mass $M_{\rm sub} = (1.9 \pm 0.1) \times 10^8 M_{\odot}$ located at (0.036 ± 0.005 arcsec, 0.576 ± 0.007 arcsec) relative to the lensing galaxy (in sky coordinates defined positive towards the west and the north, respectively), was preferred by a factor of $e^{65}$ over a smooth model. This would heuristically correspond to a $12\sigma$ detection of the substructure, if the posterior probability distribution function were Gaussian. This agrees

---

[1]Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2]Department of Physics, University of California, Davis, California 95616, USA. [3]ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands. [4]Department of Physics, University of California, Santa Barbara, California 93106, USA. [5]Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands.

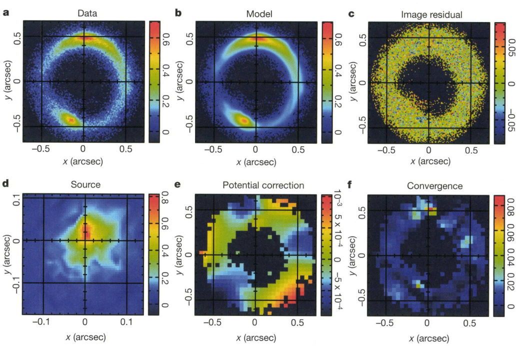


**Figure 1 | Detection of a dark-matter-dominated satellite galaxy in the gravitational lens system B1938+666 at $z = 0.881$.** The data shown here were measured at 2.2 μm using the W. M. Keck telescope in June 2010. Additional data sets at 1.6 μm, from the Keck telescope and the Hubble Space Telescope, are in Supplementary Information. **a**, Original data set with the lensing galaxy subtracted; **b**, final reconstruction; **c**, image residuals; **d**, source reconstruction; **e**, dimensionless potential correction from a smooth potential required by the model to fit the data; **f**, resulting dimensionless projected density corrections. Colour bars show counts for **a**–**d**. The total lensing potential is defined as the sum of an analytic potential for the host galaxy plus the local pixelated potential corrections defined on a Cartesian grid. The potential corrections are a general correction to the analytical smooth potential and correct for the presence of substructure, for large-scale moments in the density profile of the galaxy and for shear. When the Laplace operator is applied to the potential corrections and translated into surface density corrections, the terms related to the shear and mass sheets become zero and a constant, respectively. A strong positive density correction is found on the top part of the lensed arc. Note that these images are set on an arbitrary regular grid that has the origin shifted relative to the centre of the smooth lens model by $\Delta x = 0.024$ arcsec and $\Delta y = 0.089$ arcsec. When this shift is taken into account, the position of the density correction is consistent with the position of the substructure found in the analytic reconstruction (Supplementary Information).

well with the substructure mass found by the pixelated potential correction method described above, given the systematic and statistical uncertainties (Supplementary information).

We also considered a model containing two substructures. The corresponding Bayesian evidence was significantly lower than the Bayesian evidence of the smooth model and that of the single-substructure model. Above, we have quoted the statistical uncertainties for the substructure mass and position, but the uncertainties will also be related to several sources of systematic error, which are discussed in Supplementary Information. In general, the uncertainty in the substructure mass (under the assumption that it is located in the lens plane) is entirely dominated by the inference in the tidal truncation radius, which requires the substructure to be de-projected, yielding a systematic uncertainty of 0.45 dex.

On the basis of the detection of the substructure, we calculated the joint posterior probability function for the projected mass fraction of the halo that is made up of substructure[18], $f$, and the slope, $\alpha$, of the substructure mass function (where $\mathrm{d}N \propto M^{-\alpha}\mathrm{d}M$, where $N$ is the number density of subhaloes per comoving volume and $M$ is the substructure mass; Supplementary Information). For the JVAS B1938+666 lensing galaxy ($M_{\mathrm{lens}} = 2.46 \times 10^{10} M_{\odot}$ within an Einstein radius of 3.39 kpc), we found that $f = 3.9^{+3.6}_{-2.4}\%$ at the 68% confidence level, using a uniform prior probability distribution for $\alpha$, for a substructure mass range of $4.0 \times 10^{6} M_{\odot}$ to $4.0 \times 10^{9} M_{\odot}$. For a Gaussian prior probability distribution for $\alpha$ centred at $1.9 \pm 0.1$, which is the slope of the mass function predicted from simulations[13–15], $f = 1.5^{+1.5}_{-0.9}\%$ at the 68% confidence level. The predicted fraction of substructure from simulations within the same mass range and projected distance from the host halo centre is $\sim 0.1^{+0.3}_{-0.1}\%$ (ref. 19), which is marginally smaller than the lower limit implied by our detection of a substructure at the 68% confidence level, independent of the prior set on the slope of the mass function. However, these simulations model the formation of a Milky Way halo at $z = 0$, and simulations of elliptical galaxies out to $z \approx 1$ must be made before a more quantitative conclusion can be drawn.

Whereas flux ratio anomalies of multiply imaged quasars have previously been used to measure statistically the level of substructure in cosmologically distant lens galaxies[20–23], these analyses have degeneracies when the substructures are dark and it is difficult to localize and measure the masses of individual substructures. Hence, the shape of the substructure mass function cannot be constrained and so far it has remained unclear whether the results from quasar flux ratio anomalies are in agreement or disagreement with numerical simulations[19]. Our new low-mass detection also allows us to constrain the slope of the substructure mass function for galaxies other than our own, when combined with the detection of the ~18-fold more massive ($3.5 \times 10^{9} M_{\odot}$) substructure in the previously reported elliptical lens galaxy SDSS J0946+1006 ($M_{\mathrm{lens}} = 2.45 \times 10^{10} M_{\odot}$ within an Einstein radius of 4.60 kpc) at redshift $z = 0.222$ (ref. 8). Combining both detections resulted in a slope of $\alpha = 1.1^{+0.6}_{-0.4}$ for the mass function and an average mass fraction of $\langle f \rangle = 3.3^{+3.6}_{-1.8}\%$ for elliptical galaxies at the 68% confidence level. These results suggest that the slope of the mass function for elliptical galaxies is similar to that observed for Milky Way satellites, but that elliptical galaxies have higher mass fractions.

So far, detailed studies of the mass and luminosity properties of substructures have been limited to the Milky Way[12,24–26] and to some extent also to M31 (ref. 27), which lies at a distance of ~1 Mpc from the Milky Way. About 30 luminous satellite galaxies detected within the Milky Way virial radius (~250 kpc) are considered possible cases of cold dark matter substructure, albeit at a much lower abundance than is predicted from simulations[15]. Twenty-three of the Milky Way satellites have been found to have masses of $\sim 10^7 M_\odot$ within a 300-pc projected radius, from observations of the dynamics of their stars[28]. This method of determining masses is limited to the Local Group owing to the faintness of the satellite galaxies, and can have a systematic error if the satellite is not in dynamic equilibrium or if there is foreground contamination. The three-dimensional mass of the JVAS B1938+666 substructure within a radius of 300 pc is $M_{300} = (7.2 \pm 0.5) \times 10^7 M_\odot$ ($M_{300} = 3.4 \times 10^7 M_\odot$ for a singular isothermal sphere model), which is comparable to the masses of the Milky Way satellites[28], but is hosted by an elliptical galaxy with a velocity dispersion of $\sigma_{\rm SIS} = 187$ km s$^{-1}$ at a co-moving distance of $\sim$3 Gpc, corresponding approximately to a time when the Universe was half the age it is today. A $3\sigma$ upper limit of $L_{\rm V} < 5.4 \times 10^7 L_{\rm V,\odot}$ was found for the luminosity of the substructure in the rest-frame V band within the tidal radius, $r_{\rm t} = 440 \pm 5$ pc ($L_{\rm V,\odot}$, V-band solar luminosity). This is about a factor of four brighter than both the Sagittarius and the Fornax dwarf galaxies[12]. The velocity dispersion of the satellite, based on the Einstein radius found from the best-fitting model, is $\sigma_v \approx 16$ km s$^{-1}$, which corresponds to a circular velocity of $v_{\rm circ} \approx 27$ km s$^{-1}$. Its three-dimensional mass within 600 pc is $M_{600} = (1.13 \pm 0.06) \times 10^8 M_\odot$. The observed properties (mass and circular velocity) are comparable to those of the Sagittarius dwarf galaxy, which is a satellite of the Milky Way[12].

The mass-to-light ratios of low-mass satellites of the Milky Way have also been found to disagree with the expectations from simulations. It has been predicted that satellites with the luminosities of Fornax and Sagittarius should have a velocity dispersion that is a factor of two greater than has been observed ($\sim$16 km s$^{-1}$), which represents another problem for the cold dark matter paradigm[6]. In the case of the JVAS B1938+666 substructure, the $3\sigma$ upper limit on the luminosity is in the range of the Fornax and Sagittarius satellites, and the resulting mass-to-light ratio has a lower limit of $3.5 M_\odot / L_{\rm V,\odot}$. Although this result is consistent with the mass-to-light ratios predicted from simulations, the limit on the luminosity of the substructure will need to be decreased by an order of magnitude before any meaningful comparison can be made. This will only be possible with the next generation of large optical telescopes, which is planned to come into operation by the end of the decade.

**Received 26 June; accepted 25 October 2011.**

1. Kravtsov, A. Dark matter substructure and dwarf galactic satellites. *Adv. Astron.* **2010,** 281913 (2010).
2. Kauffmann, G., White, S. D. M. & Guiderdoni, B. The formation and evolution of galaxies within merging dark matter haloes. *Mon. Not. R. Astron. Soc.* **264,** 201–218 (1993).
3. Klypin, A., Kravtsov, A. V., Valenzuela, O. & Prada, F. Where are the missing galactic satellites? *Astrophys. J.* **522,** 82–92 (1999).
4. Moore, B. *et al.* Dark matter substructure within galactic halos. *Astrophys. J.* **524,** L19–L22 (1999).
5. Boylan-Kolchin, M., Bullock, J. S. & Kaplinghat, M. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Mon. Not. R. Astron. Soc.* **415,** L40–L44 (2011).
6. Boylan-Kolchin, M., Springel, V., White, S. D. M. & Jenkins, A. There's no place like home? Statistics of Milky Way-mass dark matter haloes. *Mon. Not. R. Astron. Soc.* **406,** 896–912 (2010).
7. Busha, M. T. *et al.* Statistics of satellite galaxies around Milky Way-like hosts. *Astrophys. J.* (submitted).
8. Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T. & Gavazzi, R. Detection of a dark substructure through gravitational imaging. *Mon. Not. R. Astron. Soc.* **408,** 1969–1981 (2010).
9. Koopmans, L. V. E. Gravitational imaging of cold dark matter substructures. *Mon. Not. R. Astron. Soc.* **363,** 1136–1144 (2005).
10. Vegetti, S. & Koopmans, L. V. E. Bayesian strong gravitational-lens modelling on adaptive grids: objective detection of mass substructure in Galaxies. *Mon. Not. R. Astron. Soc.* **392,** 945–963 (2009).
11. King, L. J. *et al.* A complete infrared Einstein ring in the gravitational lens system B1938 + 666. *Mon. Not. R. Astron. Soc.* **295,** L41–L44 (1998).
12. Strigari, L. E. *et al.* Redefining the missing satellites problem. *Astrophys. J.* **669,** 676–683 (2007).
13. Diemand, J., Kuhlen, M. & Madau, P. Dark matter substructure and gamma-ray annihilation in the Milky Way halo. *Astrophys. J.* **657,** 262–270 (2007).
14. Diemand, J., Kuhlen, M. & Madau, P. Formation and evolution of galaxy dark matter halos and their substructure. *Astrophys. J.* **667,** 859–877 (2007).
15. Springel, V. *et al.* The Aquarius Project: the subhaloes of galactic haloes. *Mon. Not. R. Astron. Soc.* **391,** 1685–1711 (2008).
16. Riechers, D. A. Molecular gas in lensed $z > 2$ quasar host galaxies and the star formation law for galaxies with luminous active galactic nuclei. *Astrophys. J.* **730,** 108–123 (2011).
17. Tonry, J. L. & Kochanek, C. S. Redshifts of the gravitational lenses MG 1131+0456 and B1938+666. *Astrophys. J.* **119,** 1078–1082 (2000).
18. Vegetti, S. & Koopmans, L. V. E. Statistics of mass substructure from strong gravitational lensing: quantifying the mass fraction and mass function. *Mon. Not. R. Astron. Soc.* **400,** 1583–1592 (2009).
19. Xu, D. D. *et al.* Effects of dark matter substructures on gravitational lensing: results from the Aquarius simulations. *Mon. Not. R. Astron. Soc.* **398,** 1235–1253 (2009).
20. Mao, S. & Schneider, P. Evidence for substructure in lens galaxies? *Mon. Not. R. Astron. Soc.* **295,** 587–594 (1998).
21. Dalal, N. & Kochanek, C. S. Direct detection of cold dark matter substructure. *Astrophys. J.* **572,** 25–33 (2002).
22. Metcalf, R. B. & Zhao, H. Flux ratios as a probe of dark substructures in quadruple-image gravitational lenses. *Astrophys. J.* **567,** L5–L8 (2002).
23. Keeton, C. R., Gaudi, B. S. & Petters, A. O. Identifying lenses with small-scale structure. I. Cusp lenses. *Astrophys. J.* **598,** 138–161 (2003).
24. Walker, M. G. *et al.* Velocity dispersion profiles of seven dwarf spheroidal galaxies. *Astrophys. J.* **667,** L53–L56 (2007).
25. Gilmore, G. *et al.* The observed properties of dark matter on small spatial scales. *Astrophys. J.* **663,** 948–959 (2007).
26. Simon, J. D. & Geha, M. The kinematics of the ultra-faint Milky Way satellites: solving the missing satellite problem. *Astrophys. J.* **670,** 313–331 (2007).
27. Mateo, M. L. Dwarf galaxies of the Local Group. *Annu. Rev. Astron. Astrophys.* **36,** 435–506 (1998).
28. Strigari, L. E. *et al.* A common mass scale for satellite galaxies of the Milky Way. *Nature* **454,** 1096–1097 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** Our results are based on observations made with the W. M. Keck Observatory and the Hubble Space Telescope. S.V. is supported by a Pappalardo Fellowship at the Massachusetts Institute of Technology, L.V.E.K. is supported (in part) through an NWO-VIDI program subsidy, and D.J.L. and C.D.F. acknowledge support from the National Science Foundation. The authors are grateful to P. Marshall for comments and feedback.

**Author Contributions** S.V. and L.V.E.K. developed the gravitational imaging technique used for the detection of substructure. S.V. carried out the gravitational lens modelling of the data with help from L.V.E.K. and J.P.M. S.V., L.V.E.K. and J.P.M. wrote the manuscript with comments from all of the authors. C.D.F. was the principal investigator of the observing programme and was responsible, along with D.J.L., for acquiring the data. D.J.L. and M.W.A. reduced the data and performed the galaxy subtraction with help from C.D.F. M.W.A. calculated the systematic error in the substructure mass. C.D.F. calculated the galaxy luminosity.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.V. (svegetti@space.mit.edu)

doi:10.1038/nature10707

# Coherent singlet-triplet oscillations in a silicon-based double quantum dot

B. M. Maune[1], M. G. Borselli[1], B. Huang[1], T. D. Ladd[1], P. W. Deelman[1], K. S. Holabird[1], A. A. Kiselev[1], I. Alvarado-Rodriguez[1], R. S. Ross[1], A. E. Schmitz[1], M. Sokolich[1], C. A. Watson[1], M. F. Gyure[1] & A. T. Hunter[1]

**Silicon is more than the dominant material in the conventional microelectronics industry: it also has potential as a host material for emerging quantum information technologies. Standard fabrication techniques already allow the isolation of single electron spins in silicon transistor-like devices. Although this is also possible in other materials, silicon-based systems have the advantage of interacting more weakly with nuclear spins. Reducing such interactions is important for the control of spin quantum bits because nuclear fluctuations limit quantum phase coherence, as seen in recent experiments in GaAs-based quantum dots[1,2]. Advances in reducing nuclear decoherence effects by means of complex control[3–5] still result in coherence times much shorter than those seen in experiments on large ensembles of impurity-bound electrons in bulk silicon crystals[6,7]. Here we report coherent control of electron spins in two coupled quantum dots in an undoped Si/SiGe heterostructure and show that this system has a nuclei-induced dephasing time of 360 nanoseconds, which is an increase by nearly two orders of magnitude over similar measurements in GaAs-based quantum dots. The degree of phase coherence observed, combined with fast, gated electrical initialization, read-out and control, should motivate future development of silicon-based quantum information processors.**

Coherent control of single, isolated quantum bits (qubits) has now been demonstrated in a large variety of physical systems[8], but not in silicon. Silicon qubits have garnered interest for over a decade in part owing to the potential to exploit techniques and infrastructure from the established microelectronics industry[9,10]. Interest in silicon is further supported by the expectation of the material's superior spin coherence times relative to other semiconductors, a result of both the reduced hyperfine coupling to nuclear spins in the semiconducting host material and the reduced number of nuclear spins. Despite the significant interest in silicon-based electrically gated qubits, progress has been hampered by several fabrication and design challenges. For example, the larger in-plane effective electron mass in silicon (which is nearly three times larger than in GaAs) shrinks the electron wavefunctions where the quantum dots are formed. Smaller wavefunctions necessitate the fabrication of correspondingly smaller devices to facilitate the isolation of a single electron in each dot. Another concern is the valley degeneracy present in bulk silicon, which is removed by an appropriate combination of strain, spatial confinement and interface properties. Considerable progress has been made in overcoming these and other obstacles, and has led to recent measurements demonstrating sufficiently large valley splitting and sufficiently long spin relaxation times for spin-qubit initialization and read-out operations[11–19].

In the present work, we form quantum dots by depleting charge from a two-dimensional electron gas within a Si/SiGe undoped heterostructure with lithographically patterned electrostatic gates (Fig. 1). We adjust the carrier concentration within the silicon quantum well using an isolated global gate above the device, and we control the charge states of the two quantum dots using the other gate electrodes. After initially depleting the double dot of all but two electrons (with one in each dot), we optimize the coupling of the quantum dots and electron-bath tunnel barriers to allow coherent dot manipulations. This and similar devices and the design advances that have facilitated coherent operation are further described in ref. 11; the work presented here used device C of that report.

As previously described[1] for a GaAs double quantum dot, the coherent electron manipulations we discuss involve operating the silicon double dot in the (1,1) charge configuration, where two electrons are separated with one in each dot. In this configuration and at a finite magnetic field, which splits off the $m = \pm 1$ spin triplet states, $(1,1)T_{\pm}$, the spin singlet state, (1,1)S, and the $m = 0$ spin triplet state, $(1,1)T_0$, may be treated as the two states of a qubit (Supplementary Information). To measure these states, we rely on spin-to-charge conversion based on Pauli spin blockade. Figure 2a shows a spin blockade signature at the (0,2)–(1,1) anticrossing obtained by cyclically pulsing the gates to transfer the double dot through the $(0,1) \rightarrow (1,1) \rightarrow (0,2)$ sequence. During the $(0,1) \rightarrow (1,1)$ transition, an electron is loaded from the bath into the left-hand dot. Both singlet and triplet states, including states formed from different valleys, are populated. Whereas the lowest-energy singlet state is able to freely transfer into (0,2), other states, including the lowest-energy triplet, are blocked by the Pauli exclusion principle[15,20–22]. This blockade is indicated by a change in the conductance of a nearby quantum point contact (QPC), which discriminates between the (1,1) and (0,2) dot charge states. The measured (0,2) signal level is used to define 100% singlet probability in subsequent experiments. The conductance change between (1,1) and (0,2), together with the known fractional time spent in the measure phase and an experimentally estimated amount of blockade relaxation during measurement, allows us directly to convert measured signal to singlet probability for all data presented here (Supplementary Information).

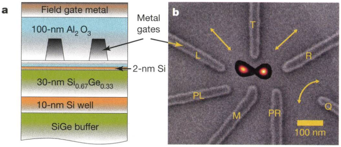


**Figure 1 | Device design.** **a**, Device cross-section showing undoped heterostructure, dielectric and gate stack. **b**, Scanning electron micrograph of actual device before dielectric isolation and field gate deposition. Electrostatic gates are labelled L (left), T, R (right), PL, M, PR and Q; gates L and R are used for fast pulsing. The straight arrows show current paths for transport experiments (Supplementary Information), whereas the curved arrow shows the path of current through the QPC. A numerical simulation of the electron density for the two-electron (1,1) state is superimposed on the micrograph.

---

[1] HRL Laboratories LLC, 3011 Malibu Canyon Road, Malibu, California 90265, USA.

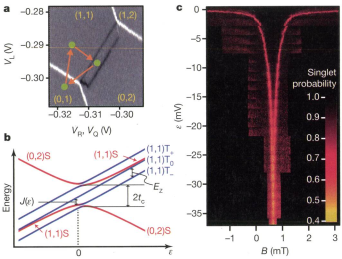


**Figure 2 | Spin blockade in the double quantum dot. a**, Subregion of the charge stability diagram, showing the differential transconductance (black and white lines) of the charge-sensing QPC as a function of gate voltages. The trapezoidal spin blockade signature is seen with cyclic pulsing through the sequence (0,1)→(1,1)→(0,2) (orange arrows). Being a differential measurement, the technique is sensitive only to changes in the average double-dot charge configuration, which is why only the edges of the trapezoidal blockade region are visible. The different colours indicate different responses of the QPC to electron movements on and between the dots. In particular, the current through the QPC decreases when electrons are added to either dot (white lines) but it increases when an electron transfers from the right dot to the left dot (black lines) (Supplementary Information). The voltage on gate Q ($V_Q$) was swept with that on gate R ($V_R$) to compensate for the capacitive coupling between gate R and the QPC and hence maintain constant sensitivity throughout the scan. **b**, Energy diagram of the (0,2)–(1,1) anticrossing, showing energies of the qubit states (1,1)S and (1,1)T$_0$ and the Zeeman split (1,1)T$_\pm$ states as functions of detuning, $\varepsilon$. The exchange-energy splitting between qubit states, $J(\varepsilon)$, the Zeeman splitting between triplet states, $E_Z$, and the tunnel coupling, $t_c$, are shown. **c**, Spin funnel obtained by measuring the degeneracy of (1,1)S and (1,1)T$_-$ as a function of $\varepsilon$ and $B$. The width of the spin funnel implies a tunnel coupling, $t_c$, of approximately 3 μeV. The functional form of $J(\varepsilon)$ extracted from the funnel data is shown in Fig. 5.

The spin blockade is lifted by the presence of an excited state, observed when biasing deeper into the (0,2) charge region. The excited state opens an additional unblocked pathway and results in the trapezoidal blockade region seen in Fig. 2a. The trapezoid height indicates an excited state at an energy of approximately 140 μeV. Numerical modelling suggests that this state is a (0,2) triplet involving an excited valley state of the right-hand dot, but it could alternatively be an orbital triplet if the right-hand dot is highly asymmetric[11,14]. In either case, the blockade region is sufficiently wide to distinguish the two qubit states (1,1)S and (1,1)T$_0$, and all other excited states are high enough in energy to have no further role in coherent operation.

The coherent evolution of the singlet and triplet qubit states is governed by two interactions. First, the singlet–triplet energy splitting, which is primarily a result of hybridization of (1,1) and (0,2) singlets, acts as an effective exchange interaction with energy $J$ between the two spins. Second, the dot-to-dot difference in the instantaneous hyperfine coupling to $^{29}$Si nuclear spins (which comprise 4.7% of the lattice nuclei in isotopically natural silicon) acts as a singlet–triplet coupling energy, $\Delta_{\mathrm{HF}}$. This coupling fluctuates from one experimental run to the next with a Gaussian distribution of standard deviation $\sigma_{\mathrm{HF}}$. In a Bloch sphere picture with the singlet and triplet states at the poles, effective exchange causes a rotation about the $z$ axis and the hyperfine energy introduces a random $x$ component to this rotation axis. Experimentally, we have direct control over only $J$ and are able to vary it by electrically pulsing the left- and right-hand gates to adjust the detuning, $\varepsilon$, between the dots (Fig. 2b). Because $J$ and $\Delta_{\mathrm{HF}}$ provide independent rotation axes and we only have direct control of $J$, exploration of the Bloch sphere requires us to achieve $J$ values both much larger than and much smaller than $\sigma_{\mathrm{HF}}$. For deep pulses into the (1,1) charge configuration, $J$ is expected to be proportional to $t_c^2/|\varepsilon|$, where $t_c$ is the interdot tunnel coupling[23]. The tunnel coupling was kept small to access $J$ values below the small $\sigma_{\mathrm{HF}}$ in silicon.

An estimate of the tunnel coupling is provided by measuring $J(\varepsilon)$ near $\varepsilon = 0$ in a 'spin funnel' experiment[24]. An applied magnetic field, $B$, splits the (1,1)T$_\pm$ states from (1,1)T$_0$ by the Zeeman energy, $E_Z = \pm g\mu_B B$, where $\mu_B$ is the Bohr magneton and $g$ is the effective electron $g$ factor in silicon. The energetic crossing of (1,1)T$_-$ with (1,1)S occurs where $J(\varepsilon) = |E_Z|$ (Fig. 2b). At this crossing, hyperfine-induced mixing of (1,1)S and (1,1)T$_-$ occurs, which is detected by the charge sensor as a spin blockade of the (1,1)T$_-$ state. The detuning dependence of the energetic crossing point, corresponding to the form of $J(\varepsilon)$ as a function of $B$, resembles a funnel whose width can be used to estimate $t_c$ (Fig. 2c). Using the funnel, we tuned our device until we obtained an acceptable $t_c$ value of about 3 μeV.

Controlled rotations between superpositions of the singlet and triplet states may be achieved by modifying $J$. Such an experiment, following ref. 1, produces data analogous to Rabi oscillations. We start with the (0,2)S ground state and then slowly separate the electrons by biasing the dots deep into the (1,1) charge configuration (Fig. 3a, b). As a result of adiabatic passage (Supplementary Information and ref. 1), a low-$J$ ground state is prepared that is a superposition of the (1,1)S and (1,1)T$_0$ states owing to the hyperfine field but, because $J$ is non-zero, is not a hyperfine eigenstate. This superposition is rotated on the Bloch sphere by the non-adiabatically applied, larger-$J$ exchange pulse. The rotation axis depends on both $J$ and $\Delta_{\mathrm{HF}}$ (Fig. 3d). After the finite-duration exchange pulse, the gates are then pulsed back deep into (1,1) and adiabatically ramped into (0,2) for read-out. During this process, the low-$J$ ground state is mapped back to the Pauli-unblocked hybridized singlet and the orthogonal low-$J$ state is mapped to the Pauli-blocked triplet.

Figure 3c shows the resulting oscillation in the singlet probability, $P_s$, as a function of the exchange-pulse duration and detuning. At less negative detunings, corresponding to high values of $J$, the initial average visibility of the Rabi oscillations is 0.7 ± 0.1. At more negative values of $\varepsilon$, corresponding to lower values of $J$, the initial visibility is reduced and hyperfine-induced damping is observed. At intermediate detunings, as many as 20 oscillations can be resolved. The average visibility, the damping at low $J$ values and the phase modulation of the oscillations closely follow the behaviour expected when averaging over a Gaussian distribution of hyperfine differences $\Delta_{\mathrm{HF}}$, as further described in Supplementary Information. Curve-fitting the data to this model allows us to extract $J(\varepsilon)$ over a wide bias range.

The practicality of implementing quantum computation using silicon-based double quantum dots depends on how long a separated (1,1)S state takes to dephase into an incoherent mixture of (1,1)S and (1,1)T$_0$ over multiple experimental runs. This timescale, $T_2^*$, directly impacts quantum control requirements, because it constrains either the speed or the complexity of high-fidelity gate operations. In III–V semiconductor quantum dots, $T_2^*$ is reported to be a few to tens of nanoseconds[1,2,5,25]. The measurement of $T_2^*$ is accomplished by preparing a state orthogonal to the hyperfine rotation axis and then allowing it to evolve freely in many trials. In each trial, the hyperfine rotation frequency is different, such that on averaging over those trials a dephasing-induced decay is observed. In our double-dot system, we prepare the initial singlet state by biasing the device into the (0,2)S state and quickly (within ~20 ns) pulsing the gates to a large negative detuning deep in the (1,1) charge configuration (Fig. 4a, b). The pulsing is non-adiabatic with respect to the nuclear mixing time, so the separated electrons are initially in the (1,1)S state. The subsequent state precession is a function of both the instantaneous hyperfine gradient and the remaining exchange energy. As a result, an ensemble average both oscillates and dephases into a mixture of (1,1)S and

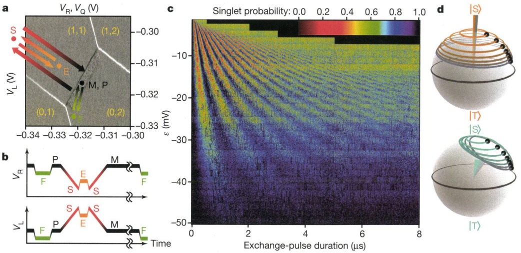


**Figure 3 | Rabi oscillation pulse sequence and data.** **a**, **b**, Essential elements of the pulse sequence used to obtain Rabi oscillations: $V_L$ and $V_R$ superimposed on the charge stability diagram (**a**); and represented as functions of time (**b**). The (0,2)S ground state is first prepared at point F with an exchange of an electron with the bath of the right-hand dot, if necessary. We then pulse to an intermediate point, P, before the subsequent adiabatic transition into the low-$J$ ground state at point S, deep into (1,1). An exchange pulse and state rotation is then applied at E and followed by a return to S. The adiabatic transfer is then reversed and the dots are biased to point M for measurement, which generally consists of >90% of the pulse cycle. **c**, Rabi oscillations of the singlet probability, $P_s$, as a function of ε and exchange-pulse duration. **d**, Bloch sphere representations of state evolution for $J > \sigma_{HF}$ (top) and $J < \sigma_{HF}$ (bottom). Several instances of initial hyperfine ground states are shown as black dots (all with $\Delta_{HF} > 0$, for clarity), with corresponding rotation axes and trajectories. The data in **c** correspond to the average probability of returning to the initial prepared low-$J$ ground state.

(1,1)$T_0$ (refs 23, 26). The gates are then pulsed quickly back into the (0,2) charge configuration for measurement.

Figure 4c shows the measured ensemble-averaged singlet probability as a function of separation time for different detunings, corresponding to different values of residual exchange energy, $J$, present during the separation pulse. For moderate negative detunings ($J > \sigma_{HF}$), the singlet state rotates primarily around the $z$ axis and the singlet probability remains close to 100% (Fig. 4d). For large negative detunings ($J < \sigma_{HF}$), the singlet state rotates around the $x$ axis and dephases, bringing the singlet probability to approximately 50%. Intermediate detunings result in oscillation frequencies and asymptotic probabilities between these two extremes. Fitting the data in Fig. 4c with a consistent model[23,26] for all traces provides estimates for the residual values of $J$ and $\sigma_{HF}$ (Supplementary Information). The extracted functional form of $J(\varepsilon)$ is shown in Fig. 5 and is in excellent agreement with the values extracted from the Rabi oscillation periods where the two data sets overlap. The model yields an estimate for $\sigma_{HF}$ of 2.6 ± 0.2 neV. A theoretical calculation of the expected $\sigma_{HF}$ value in isotopically natural silicon, using real-space, hybrid Poisson–Schrödinger/full-configuration-interaction simulation methods, and using our best estimates of the microscopic device geometry and tuning configuration, yields a value for $\sigma_{HF}$ of 1.9 neV (Supplementary Information and ref. 27), in reasonable agreement with the experimentally measured value. This value corresponds to a nuclei-induced dephasing time of $T_2^* = \sqrt{2}\hbar/\sigma_{HF} = 360 \pm 30$ ns.

The reduced hyperfine coupling in silicon relative to other material systems will allow high-fidelity quantum control of future silicon-based qubits. Although the present device has only limited coherent

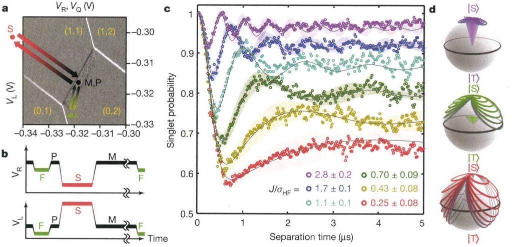


**Figure 4 | $T_2$* pulse sequence and data.** **a**, **b**, Essential elements of $T_2$* pulse sequence. Analogous with the Rabi oscillation pulse sequence (Fig. 3a, b), we initialize in the (0,2)S ground state at F and then pulse to intermediate point P. We then non-adiabatically pulse deeply into (1,1) to point S (this is the separation pulse) for a variable amount of time for singlet dephasing, and then return non-adiabatically to point M for measurement. **c**, Evolution of (1,1)S as indicated by $P_s$, as a function of separation time for several different separation-pulse detunings (residual $J$). Coloured circles show measured data; the least-squares fit (black lines) and 1-s.d. confidence intervals (coloured bands) for a hyperfine-averaged model[23,26] are superimposed. The values of $J/\sigma_{HF}$ resulting from this fit are shown in the key. **d**, Bloch sphere representations of an ensemble of trajectories, with $J/\sigma_{HF}$ values as indicated by colour in **c**. The ensemble of rotation axes is also shown in each case. The data in **c** correspond to the average projection of trajectories onto the $z$ axis.

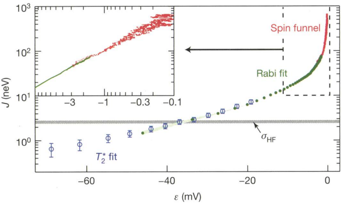


**Figure 5 | Effective exchange energy versus detuning.** $J(\varepsilon)$ extracted from the spin funnel data (Fig. 2c), Rabi oscillation data (Fig. 3c) and $T_2$* data (Fig. 4c). $J$ ranges from 0.6 to 700 neV with excellent agreement where these data sets overlap (Supplementary Information). Error bars on the $T_2$* data (blue) and the confidence interval for the Rabi data (green) are 1-s.d. fit-parameter uncertainties; spin funnel symbol sizes correspond to Gaussian-fit transition widths. The grey band denotes the estimated $\sigma_{HF}$ value from the $T_2$* data.

control capabilities owing to its reliance on the small and fluctuating hyperfine field for one axis of control, future devices will almost certainly not rely on hyperfine interactions. Other control methods, such as exchange-only control of three-dot systems[28], will allow complete coherent manipulation on a timescale much shorter than $T_2$*. There are no fundamental problems with delivering exchange pulses on subnanosecond timescales[1], which will further allow pulse compensation and dynamical decoupling techniques for high-fidelity gate operations[29,30]. These techniques can be used in future efforts to study and eliminate other decoherence sources not yet isolated in the experiments described here. Our results therefore indicate strong potential for the execution of the quantum error correction protocols that allow sustained quantum computation in silicon-based systems.

## METHODS SUMMARY

The device is characterized in a dilution refrigerator with an estimated effective electron temperature of 150 mK. The QPC conductance changes when electrons are loaded, unloaded or transferred between the dots and is used to determine the charge state of the quantum dots. Fast voltage pulses are applied to the left- and right-hand gates using a Tektronix 5014 arbitrary wave generator with low-pass filtering at its output. Measurements are conducted with an external in-plane magnetic field of 30 mT, except for the spin funnel experiment, in which the field is swept.

The differential transconductance measurement technique used is implemented by superimposing a low-frequency voltage square wave (generally 600–900 Hz) on the left-hand gate and using standard lock-in amplification techniques with a 30-ms time constant and 60-ms delay between measurements. The low-frequency square wave is achieved by concatenating an appropriate number of the higher-frequency coherent manipulation pulse sequences to bring the lock-in measurement frequency down to near the noise floor of our measurement system. The added square wave is present only during the measurement phase of the pulse sequences. Apart from a background signal resulting from coupling between the low-frequency square wave and the QPC, the differential transconductance measurement technique is sensitive only to changes in the QPC conductance resulting from changes in the double-dot charge state configuration. As a result, we see only lines marking the outlines of the charge stability regions and of the trapezoidal triplet blockade region in (0,2), instead of the more commonly seen QPC conductance plateaux.

More details on experimental methods are available in Supplementary Information.

**Received 24 June; accepted 1 November 2011.**

1. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309,** 2180–2184 (2005).
2. Koppens, F. H. L., Nowack, K. C. & Vandersypen, L. M. K. Spin echo of a single electron spin in a quantum dot. *Phys. Rev. Lett.* **100,** 236802 (2008).
3. Reilly, D. J. *et al.* Suppressing spin qubit dephasing by nuclear state preparation. *Science* **321,** 817–821 (2008).
4. Barthel, C., Medford, J., Marcus, C. M., Hanson, M. P. & Gossard, A. C. Interlaced dynamical decoupling and coherent operation of a singlet-triplet qubit. *Phys. Rev. Lett.* **105,** 266808 (2010).
5. Bluhm, H. *et al.* Dephasing time of GaAs electron-spin qubits coupled to a nuclear bath exceeding 200 μs. *Nature Phys.* **7,** 109–113 (2011).
6. Simmons, S. *et al.* Entanglement in a solid-state spin ensemble. *Nature* **470,** 69–72 (2011).
7. Tyryshkin, A. M. *et al.* Electron spin coherence exceeding seconds in high purity silicon. *Nature Mater.* (in the press); preprint at ⟨http://arxiv.org/abs/1105.3772⟩ (2011).
8. Ladd, T. D. *et al.* Quantum computers. *Nature* **464,** 45–53 (2010).
9. Kane, B. E. A silicon-based nuclear spin quantum computer. *Nature* **393,** 133–137 (1998).
10. Eriksson, M. A. *et al.* Spin-based quantum dot quantum computing in silicon. *Quantum Inf. Process.* **3,** 133–146 (2004).
11. Borselli, M. G. *et al.* Pauli spin blockade in undoped Si/SiGe two-electron double quantum dots. *Appl. Phys. Lett.* **99,** 063109 (2011).
12. Goswami, S. *et al.* Controllable valley splitting in silicon quantum devices. *Nature Phys.* **3,** 41–45 (2007).
13. Xiao, M., House, M. G. & Jiang, H. W. Parallel spin filling and energy spectroscopy in few-electron Si metal-on-semiconductor-based quantum dots. *Appl. Phys. Lett.* **97,** 032103 (2010).
14. Borselli, M. G. *et al.* Measurement of valley splitting in high-symmetry Si/SiGe quantum dots. *Appl. Phys. Lett.* **98,** 123118 (2011).
15. Lai, N. S. *et al.* Pauli spin blockade in a highly tunable silicon double quantum dot. *Sci. Rep.* **1,** 110 (2011).
16. Hayes, R. R. *et al.* Lifetime measurements (T1) of electron spins in Si/SiGe quantum dots. Preprint at ⟨http://arxiv.org/abs/0908.0173⟩ (2009).
17. Xiao, M., House, M. G. & Jiang, H. W. Measurement of the spin relaxation time of single electrons in a silicon metal-oxide-semiconductor-based quantum dot. *Phys. Rev. Lett.* **104,** 096801 (2010).
18. Morello, A. *et al.* Single-shot readout of an electron spin in silicon. *Nature* **467,** 687–691 (2010).
19. Simmons, C. B. *et al.* Tunable spin loading and $T_1$ of a silicon spin qubit measured by single-shot readout. *Phys. Rev. Lett.* **106,** 156804 (2011).
20. Johnson, A. C. *et al.* Triplet–singlet spin relaxation via nuclei in a double quantum dot. *Nature* **435,** 925–928 (2005).
21. Liu, H. W. *et al.* Pauli-spin-blockade transport through a silicon double quantum dot. *Phys. Rev. B* **77,** 073310 (2008).
22. Shaji, N. *et al.* Spin blockade and lifetime-enhanced transport in a few-electron Si/SiGe double quantum dot. *Nature Phys.* **4,** 540–544 (2008).
23. Laird, E. A. *et al.* Effect of exchange interaction on spin dephasing in a double quantum dot. *Phys. Rev. Lett.* **97,** 056801 (2006).
24. Petta, J. R. *et al.* Dynamic nuclear polarization with single electron spins. *Phys. Rev. Lett.* **100,** 067601 (2008).
25. Press, D. *et al.* Ultrafast optical spin echo in a single quantum dot. *Nature Photon.* **4,** 367–370 (2010).
26. Coish, W. A. & Loss, D. Singlet-triplet decoherence due to nuclear spins in a double quantum dot. *Phys. Rev. B* **72,** 125337 (2005).
27. Assali, L. V. C. *et al.* Hyperfine interactions in silicon quantum dots. *Phys. Rev. B* **83,** 165301 (2011).
28. DiVincenzo, D. P., Bacon, D., Kempe, J., Burkard, G. & Whaley, K. B. Universal quantum computation with the exchange interaction. *Nature* **408,** 339–342 (2000).
29. Tomita, Y., Merrill, J. T. & Brown, K. R. Multi-qubit compensation sequences. *N. J. Phys.* **12,** 015002 (2010).
30. West, J. R., Lidar, D. A., Fong, B. H. & Gyure, M. F. High fidelity quantum gates via dynamical decoupling. *Phys. Rev. Lett.* **105,** 230503 (2010).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank C. M. Marcus for discussions, J. R. Petta for assistance with measurement techniques and B. H. Fong for assistance with hyperfine calculations. Sponsored by United States Department of Defense. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the United States Department of Defense or the US Government. Approved for public release, distribution unlimited.

**Author Contributions** B.M.M., M.G.B., C.A.W., K.S.H. and A.T.H. contributed to device measurement and testing. B.H., P.W.D., I.A.-R., A.E.S. and M.S. contributed to material growth and device fabrication. B.M.M., T.D.L., A.A.K., R.S.R. and M.F.G. contributed to modelling and data analysis. B.M.M., T.D.L., A.A.K. and M.F.G prepared the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.M.M. (bmmaune@hrl.com).

doi:10.1038/nature10769

# Reconfigurable self-assembly through chiral control of interfacial tension

Thomas Gibaud[1], Edward Barry[1], Mark J. Zakhary[1], Mir Henglin[1], Andrew Ward[1], Yasheng Yang[1], Cristina Berciu[2], Rudolf Oldenbourg[3], Michael F. Hagan[1], Daniela Nicastro[2], Robert B. Meyer[1] & Zvonimir Dogic[1]

**From determining the optical properties of simple molecular crystals to establishing the preferred handedness in highly complex vertebrates, molecular chirality profoundly influences the structural, mechanical and optical properties of both synthetic and biological matter on macroscopic length scales[1,2]. In soft materials such as amphiphilic lipids and liquid crystals, the competition between local chiral interactions and global constraints imposed by the geometry of the self-assembled structures leads to frustration and the assembly of unique materials[3–6]. An example of particular interest is smectic liquid crystals, where the two-dimensional layered geometry cannot support twist and chirality is consequently expelled to the edges in a manner analogous to the expulsion of a magnetic field from superconductors[7–10]. Here we demonstrate a consequence of this geometric frustration that leads to a new design principle for the assembly of chiral molecules. Using a model system of colloidal membranes[11], we show that molecular chirality can control the interfacial tension, an important property of multi-component mixtures. This suggests an analogy between chiral twist, which is expelled to the edges of two-dimensional membranes, and amphiphilic surfactants, which are expelled to oil–water interfaces[12]. As with surfactants, chiral control of interfacial tension drives the formation of many polymorphic assemblages such as twisted ribbons with linear and circular topologies, starfish membranes, and double and triple helices. Tuning molecular chirality *in situ* allows dynamical control of line tension, which powers polymorphic transitions between various chiral structures. These findings outline a general strategy for the assembly of reconfigurable chiral materials that can easily be moved, stretched, attached to one another and transformed between multiple conformational states, thus allowing precise assembly and nanosculpting of highly dynamical and designable materials with complex topologies.**

In experiments on chiral self-assembly, we used a two-component mixture consisting of 880-nm-long rod-like fd viruses and the non-adsorbing polymer Dextran. In aqueous suspension, fd viruses alone have purely repulsive interactions[13]. Adding non-adsorbing polymer to a dilute isotropic suspension of fd rods induces attractive interactions by the depletion mechanism and leads to their condensation into colloidal membranes, which are equilibrium structures consisting of one-rod-length-thick, liquid-like monolayers of aligned rods[11] (Fig. 1a). Despite having different structures on molecular length scales, the coarse-grained properties of colloidal membranes are identical to those of conventional lipid bilayers. However, unlike their amphiphilic counterparts, colloidal membranes do not form vesicles and are instead observed as freely suspended disks with exposed edges. Here we investigate the behaviour of these exposed edges in a manner analogous to previously studied liquid–liquid domains embedded in lipid bilayers[14–16]. For our experiments, it is essential that fd viruses are chiral, that is, a pair of aligned viruses minimizes their interaction energy when they are slightly twisted with respect to each other in a preferred direction. The strength of chiral interactions can be continuously tuned to zero through either genetic or physical methods[13,17] (Supplementary Fig. 1).

Before investigating chiral membranes, we determined the edge structure of a membrane composed of simpler, achiral rods using three complementary imaging techniques, namely two-dimensional (2D) and three-dimensional (3D) polarization microscopy and electron microscopy. The local tilting of the rods within a membrane was determined using a 2D LC-PolScope, which produces images in which the intensity of each pixel represents the local retardance of the membrane[18] (Fig. 1d). Such images can be quantitatively related to the tilting of the rods away from the layer normal[19] (the *z* axis). Rods in the bulk of a membrane are aligned along the *z* axis, and it follows that 2D LC-PolScope images appear black in that region (Fig. 1e). By contrast, the bright, birefringent ring along the membrane's periphery reveals local tilting of the rods (Fig. 1e and Supplementary Fig. 2). For achiral rods, this indicates that a membrane has a hemi-toroidal curved edge (Fig. 1b, c). By comparison with an untilted edge, a curved edge structure lowers the area of the rod–polymer interface, thus reducing interfacial tension, at the cost of increasing the elastic energy due to a twist distortion. This hypothesis is confirmed by visualizing the 3D membrane structure using electron tomography, which shows that the viruses' long-axis transitions from being parallel to the *z* axis in the membrane bulk to being perpendicular to the *z* axis and tangential to the edge along the membrane periphery (Fig. 1d and Supplementary Fig. 3). For achiral viruses, the spontaneous twist at the edges is equally likely to be clockwise or anticlockwise. However, the 2D LC-PolScope records only 2D projections of the birefringence map and thus cannot distinguish between these two possibilities. For this reason, we used a 3D LC-PolScope[20], which reveals that achiral rods are equally likely to tilt in either direction, thus confirming the spontaneously broken chiral symmetry at the edge of an achiral membrane (Fig. 1f–h and Supplementary Fig. 4). Simulations of hard, achiral spherocylinders and depletant molecules provide additional verification that interfacial effects alone cause spontaneous twisting of achiral rods at the membrane's edge (Supplementary Fig. 5).

When viewed with optical microscopy, a membrane's edge shows pronounced thermal fluctuations, the analysis of which yields the line tension, $\gamma_{\mathrm{eff}}$, which is the energetic cost required to move rods from the membrane interior to the periphery[12,21,22] (Fig. 2a). A typical fluctuation spectrum for an achiral edge is shown in Fig. 2b. In the regime of small wavevector, $q$, the mean square Fourier amplitudes, $\langle a_q^2 \rangle$, are independent of $q$, allowing us to extract the effective line tension, $\gamma_{\mathrm{eff}} = k_B T / \langle a_q^2 \rangle$ (ref. 12), where $k_B$ is the Boltzmann constant and $T$ is the temperature. In the large-$q$ limit, fluctuations scale as $1/q^2$ and are therefore substantially suppressed relative to those of a simple interface. One possible explanation is that fluctuations are suppressed as a result of a quasi-one-dimensional (quasi-1D) nematic phase at the edge, which is a direct consequence of rod tilting (Fig. 1e). At large wavelengths, the elastic energy required to deform the quasi-1D nematic phase is negligible and the fluctuations are dominated by the line tension. With decreasing wavelength, there is an increasing energetic penalty associated with bending the quasi-1D nematic phase, resulting in the suppression of fluctuations.

---

[1]The Martin Fisher School of Physics, Brandeis University, 415 South Street, Waltham, Massachusetts 02454, USA. [2]Department of Biology, Brandeis University, 415 South Street, Waltham, Massachusetts 02454, USA. [3]Marine Biological Laboratory, 7 MBL Street, Woods Hole, Massachusetts 02543, USA.

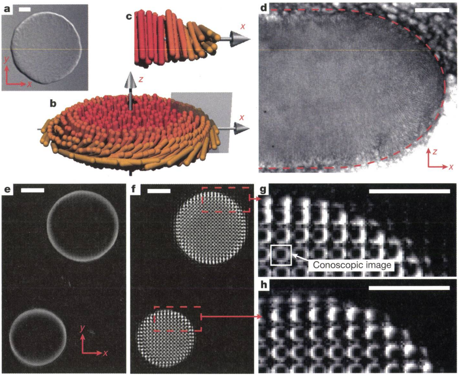


**Figure 1 | Edge structure of membranes assembled from an achiral mixture of wild-type and Tyr21Met fd phages.** **a**, Differential interference contrast (DIC) image of a colloidal membrane. **b**, Illustration of colloidal membrane indicating that its edge adopts a surface-tension-minimizing curved profile, forcing rods to twist locally. For clarity, the aspect ratio of rods has been reduced. **c**, Cross-section shown in **b**. **d**, Electron micrograph of a membrane directly visualizing the curved edge profile (Supplementary Fig. 3). **e**, 2D LC-PolScope birefringence map of two membranes. The bright band associated with the edges indicates local rod tilting (Supplementary Fig. 2). **f**, 3D LC-PolScope image composed of an array of conoscopic microimages, each sampling the local orientation of the rods. In each microimage, the displacement of the dark spot away from the centre determines the local tilting of the rods with respect to the layer normal (the *z* axis; Supplementary Fig. 4). **g**, **h**, Sections of 3D LC-PolScope images of two edges with clockwise (**g**) and anticlockwise (**h**) twists, respectively, illustrating the broken chiral symmetry found at the edge of achiral membranes. Scale bars, 5 μm (**a**, **e**–**g**, **h**); 0.2 μm (**d**).

Next we demonstrated that the chirality of the constituent rods controls the magnitude of $\gamma_{\mathrm{eff}}$. The global constraints imposed by the 2D membrane topology are fundamentally incompatible with local twisting (chiral interactions) between constituent rods, leading to the expulsion of chirality from the bulk of the membrane[7] (untwisting of rods). It follows that introducing chirality raises the energy of rods in the membrane bulk while simultaneously lowering the energy of rods close to the edge, where they are free to twist and satisfy chiral interactions. Therefore, we propose that there are two contributions to the effective line tension, $\gamma_{\mathrm{bare}}$ and $\gamma_{\mathrm{chiral}}$, such that $\gamma_{\mathrm{eff}} = \gamma_{\mathrm{bare}} - \gamma_{\mathrm{chiral}}$, where $\gamma_{\mathrm{bare}}$ is the line tension of a membrane edge composed of achiral rods and $\gamma_{\mathrm{chiral}}$ is the chiral contribution to the line tension[19,23]. To measure $\gamma_{\mathrm{bare}}$ and $\gamma_{\mathrm{chiral}}$ independently, we used the temperature dependence of fd chirality (Supplementary Fig. 1). We measured $\gamma_{\mathrm{bare}}$ at high temperature (60 °C), where rods are achiral ($\gamma_{\mathrm{chiral}} = 0$) and, therefore, $\gamma_{\mathrm{eff}} = \gamma_{\mathrm{bare}}$ (Fig. 2c). Decreasing the temperature increases the strength of chiral interactions, resulting in a smaller $\gamma_{\mathrm{eff}}$ value and demonstrating that chirality indeed reduces line tension (Fig. 2c). Temperature-dependent chiral contributions ($\gamma_{\mathrm{chiral}}(T) = -\gamma_{\mathrm{eff}}(T) + \gamma_{\mathrm{bare}}$) were converted into a function that depends on the chiral wavevector, $q_0$, using measurements of the temperature dependence of the cholesteric twist (Supplementary Fig. 1). Notably, measurements of $\gamma_{\mathrm{chiral}}$ at different depletant concentrations collapse onto a universal curve (Fig. 2d), confirming that the two contributions to $\gamma_{\mathrm{eff}}$ are uncorrelated. Our data show that chirality can reduce the line tension by as much as $450k_{\mathrm{B}}T$ per micrometre of edge length (Fig. 2d). To test this hypothesis independently, we measured $\gamma_{\mathrm{eff}}$ for membranes composed of the Tyr21Met mutant fd virus. By contrast with wild-type fd virus, the chirality of this mutant fd virus is temperature independent (Supplementary Fig. 6). We find that $\gamma_{\mathrm{eff}}$ for the Tyr21Met mutant membranes is temperature independent over the entire measurement range, unambiguously demonstrating chiral control of line tension and ruling out any intrinsic temperature effects.

Our findings raise the possibility that at sufficiently low temperatures the chiral contribution to interfacial energy could dominate the bare line tension, lowering the energetic cost of creating edges and leading to

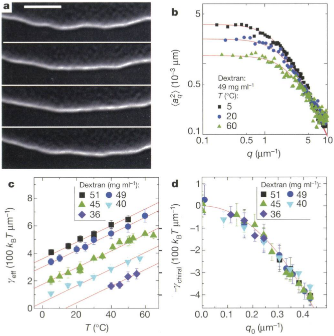


**Figure 2 | Chiral control of effective line tension, $\gamma_{\mathrm{eff}}$.** **a**, DIC images, taken 1 s apart, illustrating the fluctuations of the membrane's edge. Scale bar, 5 μm. **b**, Fluctuation spectrum ($\langle a_q^2 \rangle$ versus $q$) plotted for three different temperatures. For small $q$, $\langle a_q^2 \rangle$ is independent of $q$ and inversely proportional to the effective line tension, $\gamma_{\mathrm{eff}} = k_{\mathrm{B}}T/\langle a_q^2 \rangle$. For large $q$, the suppressed fluctuations are independent of depletant concentration (osmotic pressure) and sample temperature (fd chirality) and scale as $1/q^2$. Red lines are fits using equation (3) (Methods). **c**, $\gamma_{\mathrm{eff}}$ extracted from the low-$q$ regime, plotted for a range of sample temperatures (fd chirality) and Dextran concentrations (osmotic pressures). In the achiral limit at 60 °C, $\gamma_{\mathrm{chiral}} = 0$ and $\gamma_{\mathrm{bare}} = \gamma_{\mathrm{eff}}$. Increasing the Dextran concentration increases $\gamma_{\mathrm{bare}}$. Decreasing the temperature reduces $\gamma_{\mathrm{eff}}$ because $\gamma_{\mathrm{chiral}}$ increases. The red lines of fixed slope are guides to the eyes illustrating the universal scaling of $\gamma_{\mathrm{eff}}$ with chirality. **d**, $\gamma_{\mathrm{chiral}}$ as a function of the twist wavevector, $q_0$ (measured in the cholesteric phase), for various depletant concentrations. The temperature dependence of $q_0$ is extracted from Supplementary Fig. 1. The red curve is quadratic fit to the data: $\gamma_{\mathrm{chiral}} = Cq_0^2$. Error bars, s.d.; $n = 10$.

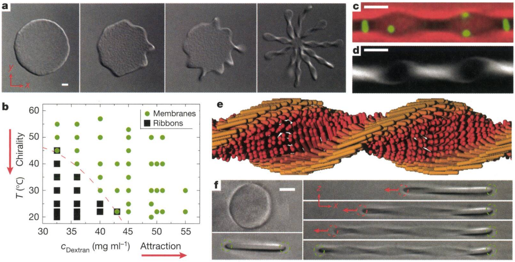


**Figure 3 | Transition of a 2D disk into 1D twisted ribbons. a,** A temperature quench reduces the line tension, inducing a transition of the 2D membrane into 1D twisted ribbons (Supplementary Movie 1). **b,** The stability diagram indicates regions of phase space where ribbons and membranes are observed as a function of temperature (chirality) and Dextran concentration ($c_{\text{Dextran}}$). The boundary indicates the temperature above which no stable or metastable ribbons are observed at a given Dextran concentration. **c,** An overlaid phase contrast (red)/fluorescence (green) image of a stable ribbon containing a low volume fraction of fluorescently labelled rods (Supplementary Movie 2). **d,** LC-PolScope image indicating rod tilting that penetrates from the edge of the ribbon towards its centre. The intensity at each pixel is proportional to $\sin^2(\theta)$, where $\theta$ is the local rod tilt angle away from the image plane. **e,** Schematic structure of twisted ribbons as deduced from optical microscopy. **f,** A 2D colloidal membrane is trapped with a dual-beam optical trap and stretched, inducing the transition to a 1D twisted ribbon (Supplementary Movie 3). Red circles indicate moving traps and green circles indicate fixed traps. Scale bars, 2 μm.

their spontaneous formation. In this respect, we expect chirality to have a role similar to that of amphiphilic surfactants in oil–water mixtures. Such surfactants lower surface tension and lead to the formation of excess interfaces through the assembly of microemulsions. To investigate this possibility, we quenched a membrane assembled at low depletant (Dextran) concentration from the high-temperature achiral phase. With decreasing temperature, fluctuations of the membrane edge become more pronounced, indicating that $\gamma_{\text{eff}}$ is decreasing. Eventually the edge becomes unstable, resulting in a remarkable polymorphic transition in which twisted, ribbon-like structures grow along the entire periphery of the disk, generating a starfish-shaped membrane (Fig. 3a and Supplementary Movie 1). This polymorphic transition is reversible; with increasing temperature, the starfish membrane collapses back into a disk. The stability diagram indicates that 1D twisted ribbons are stable at low Dextran concentrations and low temperatures (small $\gamma_{\text{bare}}$ values and large $\gamma_{\text{chiral}}$ values), whereas flat 2D membranes are stable for large $\gamma_{\text{bare}}$ values or small $\gamma_{\text{chiral}}$ values (Fig. 3b). DIC microscopy reveals the overall structure of the ribbons, such as their pitch and width (Fig. 3a); fluorescence microscopy shows the liquid-like dynamics of single rods within the assemblage (Fig. 3c and Supplementary Movie 2); and 2D LC-PolScope imaging quantitatively determines the local tilt of the rods (Fig. 3d). A schematic representation of the twisted ribbons based on this information, indicating inhomogeneous tilting of the rods within a ribbon, is shown in Fig. 3e. Such data are in agreement with recent theoretical predictions[24].

Chiral self-assembly produces structures that combine seemingly divergent properties of high elasticity and fluidity. To demonstrate these properties, we trapped two opposite sides of a flat disk with a dual-beam optical trap and applied an extensional force, causing the transition from a stretched disk into a twisted ribbon (Fig. 3f and Supplementary Movie 3). This mechanically induced disk-to-ribbon transition is reversible; on removal of the optical trap, the highly elastic ribbon relaxes back into its original shape. Furthermore, using optical tweezers it is possible to change the architecture of the chiral assemblages systematically, allowing new pathways for assembling nanomaterials with highly complex topologies. For example, two ends of a linear twisted ribbon are easily joined together to assemble a closed, ring-like polymeric structure. The disk-to-ribbon transition is tightly coupled to the assemblage topology (Fig. 4d). On increasing the temperature, all linear ribbons transform into achiral disks (Supplementary Movie 4), whereas closed, ring-like ribbons remain twisted owing to the constraint imposed by the ring topology (Supplementary Movie 5).

The chiral assembly pathways described here are hierarchical, opening up the possibility that simple changes on microscopic (ångström) length scales can be used to control structures on macroscopic (millimetre) length scales[25,26]. At the smallest relevant length scale, each virus is assembled from DNA that is coated with thousands of

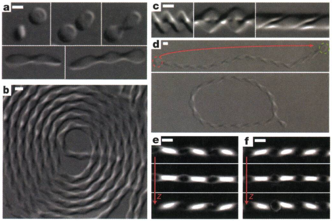


**Figure 4 | Hierarchical self-assembly: from isolated viruses and metastable disks to singly and doubly twisted ribbons. a,** In the ribbon region of the phase diagram in Fig. 3b, fd viruses condense into metastable, self-limited disks. On longer timescales, these disks coalesce and form 1D twisted ribbons. **b,** Image of a single ribbon that has collapsed into a toroid. **c,** Doubly twisted ribbons are assembled by wrapping around each other two singly twisted ribbons with a well-defined phase difference (Supplementary Fig. 8). **d,** Directed assembly of a ring-like supramolecular polymer from a linear twisted ribbon using optical tweezers. **e, f,** Single point mutation of the major coat protein permits microscopic control of ribbon chirality. Wild-type fd virus assembles into left-handed ribbons (**e**), whereas the Tyr21Met fd virus forms right-handed ribbons (**f**), as demonstrated by the *z*-stack series of 2D LC-PolScope images. Scale bars, 2 μm.

copies of the major coat protein. At the mesoscopic length scale, attractive depletion interactions condense dilute, rod-like viruses into self-limited, metastable, mesoscopic colloidal disks whose diameter is approximately one twist penetration length[19,23,27,28] (Fig. 4a). Once assembled, mesoscopic disks themselves experience mutual attractions, resulting in lateral coalescence and formation of twisted doublets. On longer timescales, this coalescence process continues, producing ribbons with complex topologies and multiple branching points (Supplementary Fig. 7). The length of these ribbons can reach macroscopic (millimetre) dimensions. This hierarchical assembly process describes a robust synthesis of ribbons that behave as supramolecular polymers and can assemble into even more-complex structures at higher levels of hierarchy, such as toroids (Fig. 4b). Under certain conditions, ribbons also wrap around each other to produce a great variety of double- and triple-helical structures (Fig. 4c and Supplementary Fig. 8). A challenge associated with hierarchical assembly pathways is to control the final macroscopic assemblage by specific modification of relevant microscopic parameters. Previous work showed that even a single amino-acid mutation of the major coat protein can alter the microscopic chirality of individual viruses[13]. We used this finding to demonstrate that such amino-acid exchange propagates across all hierarchical levels, switching the ribbon handedness and allowing for microscopic control of macroscopic ribbon chirality (Fig. 4e, f).

We have shown that the chirality of the constituent molecules can be used to tune the line tension associated with the exposed edge of a colloidal monolayer membrane. The potential of hierarchical chiral self-assembly is demonstrated by assembling reconfigurable materials with unique mechanical properties and complex topologies that can easily switch between multiple polymorphic forms[29,30].

## METHODS SUMMARY

Wild-type and Tyr21Met fd viruses were purified according to previously published protocols[13]. To reduce the number of longer viruses, which destabilize monolayer structures such as colloidal membranes or twisted ribbons, virus suspensions were cycled through isotropic–nematic phase coexistence. Only the isotropic portion of the samples, enriched in shorter viruses, was used for experiments[11]. The polydispersity of viruses was quantified using gel electrophoresis (Supplementary Fig. 9). All samples were prepared in a buffer containing 20 mM Tris and 100 mM NaCl at pH 8.0. Addition of the non-adsorbing polymer Dextran (relative molecular mass, 500,000; Sigma-Aldrich) induces condensation of rod-like viruses into either flat colloidal membranes or various other chiral assemblages. The strength of chiral interactions between rods was tuned either by changing the sample temperature or by varying the ratio of wild-type fd virus to Tyr21Met fd virus in membranes containing a mixture of the two[13]. Optical microscopy data was taken with an inverted microscope (Nikon TE-2000) equipped with DIC optics, a homemade heating/cooling stage and a cooled charged-coupled-device camera (QImaging, Retiga Exi). The 2D and 3D LC-PolScope imaging set-ups have been described elsewhere[18,20]. To prevent binding of viruses to surfaces, microscope slides and coverslips were cleaned and subsequently coated with a silane agent (3-(trimethoxysilyl) propyl methacrylate, Sigma Aldrich) to which a polyacrylamide brush was polymerized from methacrylate groups.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 31 July; accepted 6 December 2011.**
**Published online 4 January 2012.**

1. Pasteur, L. On the relations that can exist between crystalline form, and chemical composition and the sense of rotary polarization. *Ann. Chim. Phys.* **24,** 442–459 (1848).
2. Hirokawa, N., Tanaka, Y., Okada, Y. & Takeda, S. Nodal flow and the generation of left-right asymmetry. *Cell* **125,** 33–45 (2006).
3. Harris, A. B., Kamien, R. D. & Lubensky, T. C. Molecular chirality and chiral parameters. *Rev. Mod. Phys.* **71,** 1745–1757 (1999).
4. Oda, R., Huc, I., Schmutz, M., Candau, S. J. & MacKintosh, F. C. Tuning bilayer twist using chiral counterions. *Nature* **399,** 566–569 (1999).
5. Aggeli, A. *et al.* Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta-sheet tapes, ribbons, fibrils, and fibers. *Proc. Natl Acad. Sci. USA* **98,** 11857–11862 (2001).
6. Kamien, R. D. & Selinger, J. V. Order and frustration in chiral liquid crystals. *J. Phys. Condens. Matter* **13,** R1–R22 (2001).
7. de Gennes, P. G. An anology between superconductors and smectics A. *Solid State Commun.* **88,** 1039–1042 (1993).
8. Renn, S. R. & Lubensky, T. C. Abrikosov dislocation lattice in a model of the cholesteric to smectic-A transition. *Phys. Rev. A* **38,** 2132–2147 (1988).
9. Matsumoto, E. A., Alexander, G. P. & Kamien, R. D. Helical nanofilaments and the high chirality limit of smectics *A. Phys. Rev. Lett.* **103,** 257804 (2009).
10. Hough, L. E. *et al.* Helical nanofilament phases. *Science* **325,** 456–460 (2009).
11. Barry, E. & Dogic, Z. Entropy driven self-assembly of nonamphiphilic colloidal membranes. *Proc. Natl Acad. Sci. USA* **107,** 10348–10353 (2010).
12. Safran, S. *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes* 79–85 (Addison Wesley, 1994).
13. Barry, E., Beller, D. & Dogic, Z. A model liquid crystalline system based on rodlike viruses with variable chirality and persistence length. *Soft Matter* **5,** 2563–2570 (2009).
14. Baumgart, T., Hess, S. T. & Webb, W. W. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature* **425,** 821–824 (2003).
15. Honerkamp-Smith, A. R. *et al.* Line tensions, correlation lengths, and critical exponents in lipid membranes near critical points. *Biophys. J.* **95,** 236–246 (2008).
16. Lee, K. Y. C. & McConnell, H. M. Quantitized symmetry of liquid monolayer domains. *J. Phys. Chem.* **97,** 9532–9539 (1993).
17. Dogic, Z. & Fraden, S. Cholesteric phase in virus suspensions. *Langmuir* **16,** 7820–7824 (2000).
18. Oldenbourg, R. & Mei, G. New polarized light microscope with precision universal compensator. *J. Microsc.* **180,** 140–147 (1995).
19. Barry, E., Dogic, Z., Meyer, R. B., Pelcovits, R. A. & Oldenbourg, R. Direct measurement of the twist penetration length in a single smectic a layer of colloidal virus particles. *J. Phys. Chem. B* **113,** 3910–3913 (2009).
20. Oldenbourg, R. Polarized light field microscopy: an analytical method using a microlens array to simultaneously capture both conoscopic and orthoscopic views of birefringent objects. *J. Microsc.* **231,** 419–432 (2008).
21. Aarts, D. G. A. L., Schmidt, M. & Lekkerkerker, H. N. W. Direct visual observation of thermal capillary waves. *Science* **304,** 847–850 (2004).
22. Fradin, C. *et al.* Reduction in the surface energy of liquid interfaces at short length scales. *Nature* **403,** 871–874 (2000).
23. Pelcovits, R. A. & Meyer, R. B. Twist penetration in single-layer smectic A discs of colloidal virus particles. *Liq. Cryst.* **36,** 1157–1160 (2009).
24. Kaplan, C. N., Tu, H., Pelcovits, R. A. & Meyer, R. B. Theory of depletion-induced phase transition from chiral smectic-A twisted ribbons to semi-infinite flat membranes. *Phys. Rev. E* **82,** 021701 (2010).
25. Srivastava, S. *et al.* Light-controlled self-assembly of semiconductor nanoparticles into twisted ribbons. *Science* **327,** 1355–1359 (2010).
26. Chung, W. J. *et al.* Biomimetic self-templating supramolecular structures. *Nature* **478,** 364–368 (2011).
27. Grason, G. M. & Bruinsma, R. F. Chirality and equilibrium biopolymer bundles. *Phys. Rev. Lett.* **99,** 098101 (2007).
28. Claessens, M., Semmrich, C., Ramos, L. & Bausch, A. R. Helical twist controls the thickness of F-actin bundles. *Proc. Natl Acad. Sci. USA* **105,** 8819–8822 (2008).
29. Nguyen, T. D. & Glotzer, S. C. Switchable helical structures formed by the hierarchical self-assembly of laterally tethered nanorods. *Small* **5,** 2092–2098 (2009).
30. Nguyen, T. D. & Glotzer, S. C. Reconfigurable assemblies of shape-changing nanorods. *ACS Nano* **4,** 2585–2594 (2010).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This work was supported by the US National Science Foundation (NSF-MRSEC-0820492, NSF-DMR-0955776, NSF-MRI-0923057, NSF-CMMI-1068566) and the Petroleum Research Fund (ACS-PRF 50558-DNI7). We acknowledge use of the Brandeis MRSEC optical microscopy facility.

**Author Contributions** T.G., E.B., M.J.Z., R.B.M. and Z.D. designed the experiments and interpreted the results. T.G., E.B. and M.J.Z. performed the experiments. A.W. performed the optical trapping experiments. E.B., C.B. and D.N. performed the electron microscopy imaging. M.H. performed the experiments on mutant viruses. R.O. performed the LC-PolScope imaging. Y.Y. and M.F.H. designed and performed the computer simulations. T.G., E.B., M.J.Z. and Z.D. wrote the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Z.D. (zdogic@brandeis.edu).

doi:10.1038/nature10740

# Kimberlite ascent by assimilation-fuelled buoyancy

James K. Russell[1], Lucy A. Porritt[1], Yan Lavallée[2] & Donald B. Dingwell[2]

**Kimberlite magmas have the deepest origin of all terrestrial magmas and are exclusively associated with cratons[1–3]. During ascent, they travel through about 150 kilometres of cratonic mantle lithosphere and entrain seemingly prohibitive loads (more than 25 per cent by volume) of mantle-derived xenoliths and xenocrysts (including diamond)[4,5]. Kimberlite magmas also reputedly have higher ascent rates[6–9] than other xenolith-bearing magmas[10,11]. Exsolution of dissolved volatiles (carbon dioxide and water) is thought to be essential to provide sufficient buoyancy for the rapid ascent of these dense, crystal-rich magmas. The cause and nature of such exsolution, however, remains elusive and is rarely specified[6,9]. Here we use a series of high-temperature experiments to demonstrate a mechanism for the spontaneous, efficient and continuous production of this volatile phase. This mechanism requires parental melts of kimberlite to originate as carbonatite-like melts. In transit through the mantle lithosphere, these silica-undersaturated melts assimilate mantle minerals, especially orthopyroxene, driving the melt to more silicic compositions, and causing a marked drop in carbon dioxide solubility. The solubility drop manifests itself immediately in a continuous and vigorous exsolution of a fluid phase, thereby reducing magma density, increasing buoyancy, and driving the rapid and accelerating ascent of the increasingly kimberlitic magma. Our model provides an explanation for continuous ascent of magmas laden with high volumes of dense mantle cargo, an explanation for the chemical diversity of kimberlite, and a connection between kimberlites and cratons.**

Our current understanding of kimberlite ascent is greatly hampered by uncertainty about the compositions of primary kimberlite melts[1,3,5,12]. These compositions remains elusive for three reasons: kimberlites contain abundant (>30%) xenocrystic material; kimberlite melt quenched to glass has not been observed; and kimberlites are highly susceptible to alteration owing to their ultrabasic composition[1] and age[2]. Many strategies have been used to estimate kimberlite melt compositions[12,13] but the range of postulated compositions remains large and contentious ($SiO_2 < 25$–35%, MgO 15–25%)—especially the content of dissolved volatiles ($CO_2$ and $H_2O$)[1,3,12,13]. Physical properties estimated for the putative range of melt compositions include densities similar to basaltic or komatiitic melts (2,800–2,900 $kg\,m^{-3}$) and low viscosities ($10^{-2}$–$10^{2}$ Pa s)[12].

Mantle xenoliths transported by kimberlite constrain the formation depths of these magmas to the base of the cratonic mantle lithosphere (CML) or deeper[4]. Mechanical disaggregation of xenoliths produces the mantle-olivine-dominated (>25 vol.%) suite of xenocrysts characteristic of most kimberlites (Fig. 1a). Notably, orthopyroxene (opx), which comprises ~15–27% of the CML[14–16] and is the most silicic mineral present (Fig. 1b), is rare-to-absent in kimberlite. Where observed, opx texturally records severe disequilibrium (Fig. 1c)[5,17,18]. The high crystal contents suggest bulk densities of the kimberlite magma (melt+crystal) in the range of 3.0–3.1 $g\,cm^{-3}$. Buoyant ascent of kimberlite magma from depths >120 km can be greatly enhanced by the presence of a low-density exsolved fluid phase (density of $CO_2$ fluid $\rho_{CO_2} \approx 1.2\,g\,cm^{-3}$ at pressure $P \approx 2$ GPa).

Here we present a mechanism for the rapid ascent of these deep-seated, high-density magmas. Our model is inspired by new

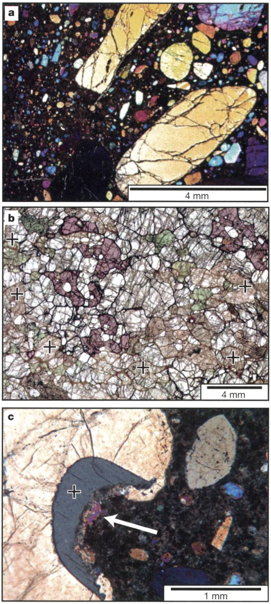


**Figure 1 | Transmitted-light optical microscope images of thin sections of kimberlite, peridotite and orthopyroxene.** **a**, Kimberlite hosts abundant rounded centimetre- to millimetre-scale mantle-derived xenocrysts of olivine (ol) and subordinate amounts of xenocrystic clinopyroxene (cpx), garnet (gar) and ilmenite (ilm). Orthopyroxene (opx) is conspicuously absent. **b**, Peridotitic xenolith from kimberlite, showing original mantle-equilibrated mineralogy (ol > opx > cpx > gar), textures, and grain size and shape distributions of mantle minerals. Grains of opx are marked with '+' (image courtesy of M. Kopylova). **c**, Kimberlite containing a fragment of disaggregated mantle peridotite (arrow), comprising a grain of partially dissolved opx (+; grey grain) within larger ol grain (yellow grain) (image courtesy of C. Brett).

---

[1] Volcanology and Petrology Laboratory, Earth and Ocean Sciences, University of British Columbia, Vancouver V6T 1Z4, Canada. [2] Department of Earth and Environmental Sciences, Ludwig Maximilian University of Munich, München 80333, Germany.

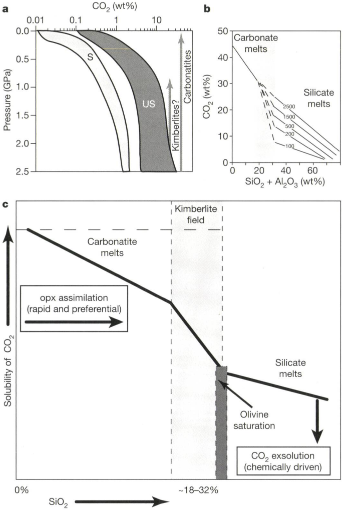


**Figure 2 | CO$_2$ solubilities in silicic to carbonate melts (after ref. 19).** **a**, CO$_2$ solubility limits for silica-saturated (S) and undersaturated (US) melts and hypothetical solubilities of carbonatite and kimberlite melt. **b**, Pressure and composition dependence of CO$_2$ solubility across the carbonate–silicate transition. Note limited effects of pressure (numbers on lines, MPa) on CO$_2$ solubility versus the strong effects of composition (that is, SiO$_2$ + Al$_2$O$_3$ wt%). **c**, Schematic model for assimilation-induced fluid-exsolution (that is, vesiculation) of carbonatitic or proto-kimberlitic melts. Opx assimilation drives non-silicate melts (left side) to more silicic compositions (towards right side), causing a decrease in CO$_2$ solubility and exsolution of a fluid phase. Continued assimilation causes continued exsolution, and results in a kimberlitic melt and olivine saturation at about 18–32 wt% SiO$_2$.

experimental data on the solubility of volatiles (that is, CO$_2$) in melts across the silicate–carbonate transition[19] (Fig. 2). This work elucidates the pressure and compositional controls on CO$_2$ solubility in silicic (for example, basalt to melilite) to carbonatitic melts (Fig. 2a,b) and provides three main insights. First, the solubility of CO$_2$ (and H$_2$O) in silicic melts increases with pressure but, importantly, in magmas with SiO$_2$ + Al$_2$O$_3$ > 35 wt%, solubility is limited to ~10–15%. This precludes extraordinary amounts of volatile being sequestered in parental kimberlitic melts[9,20]. Second, carbonatitic melts (that is, SiO$_2$ + Al$_2$O$_3$ < 20 wt%) have substantially higher CO$_2$ solubilities that are essentially independent of pressure (Fig. 2b). Solubility of CO$_2$ in carbonate melts is limited only by melt stoichiometry, and can be >40% (refs 21, 22). Last, the transition from carbonatitic to silicic (that is, kimberlitic) melt compositions is accompanied by a pronounced decrease in CO$_2$ solubility (Fig. 2b).

We suggest that melts parental to kimberlite originate as carbonatitic or near-carbonatitic melts. Such melts begin with near-stoichiometric CO$_2$ contents (~40 wt%) and are buoyant relative to the cratonic mantle lithosphere (density $\rho \approx 2.6$–$2.8\,\mathrm{g\,cm^{-3}}$)[22]. Carbonatitic melts can also accommodate substantial amounts of water (10 wt% at 0.1 GPa) without reducing their CO$_2$ contents[23]. The melts entering the CML initiate the mechanical incorporation and disaggregation of xenoliths[6,24], thereby ensuring that newly liberated xenocrysts are continuously fed into the melt. At these pressures ($P < 2.5$ GPa) and temperatures ($T > 1{,}250\,^{\circ}\mathrm{C}$), all mantle minerals are out of equilibrium and react with the low-$a_{\mathrm{SiO_2}}$ carbonatitic melt[25] ($a_{\mathrm{SiO_2}}$ is the activity of SiO$_2$). Orthopyroxene, as the most silica-saturated phase, has the highest affinity for dissolution and is assimilated rapidly and preferentially over other phases[5,17,25–27].

Assimilation of opx drives the melt towards more silicic compositions (Fig. 2c). The increased SiO$_2$ content of the contaminated carbonatitic melt causes a decrease in CO$_2$ solubility that is expressed by immediate fluid exsolution deep within the CML. Continued opx assimilation is attended by continuous fluid exsolution, leading to ever-increasing buoyancy. The process is gradual until opx assimilation produces a 'kimberlitic' melt composition (SiO$_2$ > 18 wt%), which initiates a catastrophic drop in CO$_2$ solubility. This results in massive exsolution of a volatile phase that reduces magma density, increases buoyancy, and supports rapid and accelerating ascent.

Results from a series of high-temperature melting experiments test these inductive ideas, and illustrate the potential speed and efficiency of the process. Our experiments involve melting powdered mixtures of Na$_2$CO$_3$ and natural mantle-derived opx (MgSiO$_3$; Fig. 3a), and measuring the weight loss as a proxy for CO$_2$ loss. The procedures and resulting data are fully described in Methods and Supplementary Table 1.

In non-steady-state ('transient') experiments, the melt rapidly assimilates the opx, creating a more silicic melt that is oversaturated in dissolved CO$_2$. The melt exsolves dissolved CO$_2$ (that is, undergoes decarbonation), causing a weight loss proportional to the opx content (Fig. 3b). These experiments demonstrate the spontaneous nature and rapid (~20 min) rate of the assimilation and decarbonation reactions. Reaction (that is, weight loss) continues until all of the opx has dissolved, the melt has reached its maximum SiO$_2$ content, and the CO$_2$ solubility limit of the new more silicic melt is reached (Fig. 3b). Static experiments track the change in melt composition and the production of CO$_2$ via melt decarbonation (Fig. 3c,d). The SiO$_2$ and MgO contents of the hybrid melts are greater than the bulk compositions (Fig. 3a) owing to enrichment by CO$_2$ loss. Na$_2$O content shows an initial increase and then decreases owing to the combined effects of dilution (addition of opx) and enrichment (decarbonation). CO$_2$ decreases dramatically. The addition of ~42 wt% opx produces SiO$_2$-rich (~32.7 wt%) melts in which all CO$_2$ is lost due to decarbonation (Fig. 3d). The maximum amount of CO$_2$ that would be released from the magma during assimilation to fuel the ascent of kimberlite is ~25–30 wt% (Fig. 3c, d).

Primary mantle-derived carbonatitic to near-carbonatitic (that is, low-$a_{\mathrm{SiO_2}}$) melts provide an appealing model for parental kimberlite magma. Carbonatite melts are accepted as a common product of partial melting of carbonated mantle sources at pressures >2.5 GPa (refs 21, 28). Recent experimental work shows that the solidus melts of carbonated peridotite at >2.5 GPa will always be CO$_2$ rich and SiO$_2$ poor as long as carbonate is stable in the mantle assemblage[21]. These experimentally produced melts contain in excess of 40% dissolved CO$_2$, and can even accommodate substantial dissolved H$_2$O (ref. 23); we suggest that the asthenospheric production of such melts marks the onset of kimberlite ascent (Fig. 4a, b). Our mechanistic model for kimberlite ascent (Fig. 4b, c) supports petrogenetic models that interpret the diversity of kimberlite bulk rock compositions as mechanical mixing of mantle olivine (70–80%) and a carbonate-rich melt[29] combined with assimilation of mantle orthopyroxene (20–30%)[30].

We modelled the chemical evolution of the ascending melt by assuming that the amount of opx assimilated is linearly related to the

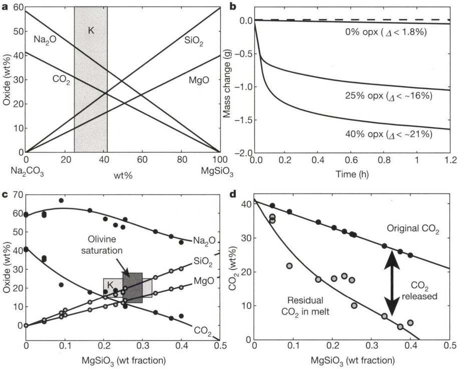


**Figure 3 | High-temperature analogue experiments.** **a**, Compositions produced by mechanical mixing of sodium carbonate ($Na_2CO_3$) and orthopyroxene (opx, $MgSiO_3$). (Data given in Supplementary Table 1). Shaded field (K) denotes range of $SiO_2$ contents for kimberlite melts based on Igwisi Hills kimberlite (Supplementary Table 2). **b**, Results of transient high-temperature experiments, showing mass change with time due to decarbonation driven by opx assimilation. Total weight loss ($\Delta$) is reported in parentheses. **c**, Melt compositions produced by chemical reaction during melting of $Na_2CO_3$-opx mixtures versus opx content. Shaded fields mark kimberlite (K) melts and condition for olivine saturation and crystallization. **d**, Post-experiment distribution of $CO_2$ between hybrid melt and the exsolved fluid phase. Distance between two lines is total exsolved (released) $CO_2$.

distance travelled in the mantle lithosphere, and that assimilation and decarbonation are essentially instantaneous (left dashed line, Fig. 4d; Supplementary Table 2). This process can be sustained energetically by the adiabatic decompression of the magma during its ascent, which can result in considerable (150–200 °C) superheat[31]. After assimilation of ~15–20% opx, a carbonatite melt will evolve to a composition close to melt compositions estimated for the Igwisi Hills kimberlite lavas[32] ($SiO_2$ 16–22 wt%; Supplementary Table 2; Fig. 4d). The amount of $CO_2$ released during assimilation and available to fuel the ascent of the contaminated carbonatite (kimberlite in the widest sense) is ~25%. Ultimately, continued opx assimilation drives the melt towards olivine saturation (Fig. 2c), allowing for late-stage magmatic crystallization of olivine as overgrowths observed on rounded olivine xenocrysts or as microphenocrysts[5,18,25,26] (Fig. 2c). Olivine crystallization would logically coincide with the later stages of kimberlite ascent and emplacement, and clearly follows on from an earlier and more extensive period of olivine disequilibrium.

The assimilation of opx is the key to this process, because it dissolves on timescales that compete with kimberlite ascent rates[25–27]. The rapid and efficient dissolution of opx, relative to other mantle xenocrysts, explains the absence of xenocrystic opx in kimberlite[5,17]. However, the main mechanistic consequence of opx assimilation is to decrease the $CO_2$ solubility of the melt, thereby driving deep-seated exsolution and effervescence of a $CO_2$-rich volatile phase (Fig. 4c). The exsolved fluid enhances the buoyancy of the ascending magma within the deep mantle, and provides an explanation for its rapid ascent and its ability to continuously entrain and transport mantle cargo to the Earth's surface.

Our model for kimberlite ascent is unique in that it begins with a low-$a_{SiO_2}$ melt containing stoichiometric amounts of dissolved $CO_2$. The $CO_2$ is stable within the melt across a wide range of pressure–temperature conditions and, thus, exsolution of the fluid phase is not driven by depressurization, as it is in most magma ascent models. Rather, $CO_2$ fluid exsolution is driven by an unavoidable chemical reaction (that is, assimilation) that creates a new more silicic melt whose $CO_2$ solubility is exceeded; the exsolution is spontaneous and immediate. Although the initial fluid production is not driven by a change in pressure, the buoyancy of the fluid-saturated magma will be greatly enhanced by the volumetric expansion of the fluids attending depressurization. This expansion allows for continuous acceleration of the magma, entrainment of greater quantities of mantle material, and decoupling (that is, separation) of a $CO_2$-rich fluid phase from the ascending magma to create precursory fenitization (alkali metasomatism) events[33].

Our model provides an explanation for the linkage between kimberlites and cratons; it is clear that, if the parental magmas are carbonatitic, that a ready source of highly reactive opx is necessary to produce a more silica-rich melt (that is, kimberlite) that will exsolve fluid in the deep mantle. The cratonic mantle lithosphere has two attributes that support this transformation: first, it is enriched in opx (15–30%) relative to other mantle lithosphere[14–16]; and second, it is thick (90–120 km), and thus provides ample opportunity for sampling of opx-rich mantle. As long as the evolving carbonatitic melt is exposed to new opx, the assimilation-induced exsolution of $CO_2$ will continue and buoyancy is maintained or increased. Therefore, we might expect the diversity of kimberlite at the Earth's surface to actually reflect the mineralogical composition, thermal state and the thickness of the underlying mantle lithosphere rather than the source region of the melt.

The corollary is that if opx were absent, this process would not be viable, because the dissolution rates for other mantle silicates (which could also promote an assimilation-driven exsolution of fluid) are too low to keep pace with ascent rate. In addition, the enthalpy of opx dissolution is less than the heat of crystallization for olivine, implying that opx assimilation can be balanced energetically by smaller quantities of olivine crystallization and a little cooling[18]. Dissolution of olivine,

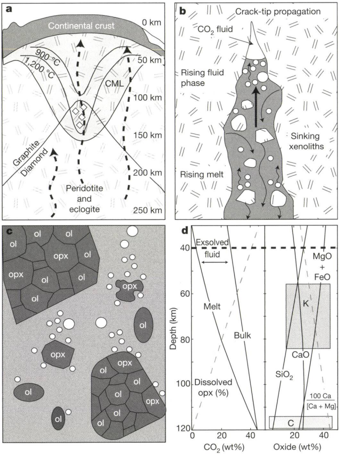


**Figure 4 | Model for assimilation-fuelled ascent of kimberlite. a**, Diverse ascent paths through cratonic mantle lithosphere (CML) shown as dashed arrows. Also shown is the line below which diamond is stable relative to graphite. **b**, Melts produced by melting of carbonated peridotite transit mantle lithosphere as dykes by crack-tip propagation, liberating dense (sinking) xenoliths to the $CO_2$-rich silica-undersaturated melt, and causing effervescence of buoyant (rising) $CO_2$-fluid. **c**, Xenoliths disaggregate and release individual mineral grains (for example, ol) to carbonatitic melt; opx grains are assimilated, preferentially promoting volatile exsolution. Deep-seated volatile production supports continued, crack-propagation-limited magma ascent. **d**, Chemical evolution of melt during ascent, including: left panel, wt% of opx assimilated (dashed line, dissolved), and $CO_2$ content in bulk system and in melt (solid lines); right panel, oxide ($SiO_2$, MgO + FeO, CaO) contents of the evolving melt (solid lines), and the 100 Ca/(Ca + Mg) ratio of melt (dashed line). Initial carbonatitic and derivative kimberlite melts (shaded boxes) defined by $SiO_2$ contents (Supplementary Table 2).

for example, would require substantial cooling of the magma (as well as longer times). In parts of the Earth where carbonatitic magmas do not transit the cratonic mantle lithosphere, there are only two possible outcomes. The carbonatitic melt successfully travels through a relatively thin lithosphere to be emplaced or erupted as carbonatite, or the carbonatite melt incorporates opx-poor mantle and crystallizes at depth because it fails to assimilate mantle silicates efficiently enough to promote a deep-seated fluid phase and the requisite buoyancy.

We suggest the possibility that all kimberlites start life as carbonatitic melts produced through partial melting of carbonated peridotite in the deep subcratonic mantle. As the melts transit the mantle lithosphere they continuously incorporate and disaggregate peridotite xenoliths. The low $a_{SiO_2}$ of the carbonatitic melt causes the rapid and preferential dissolution of opx relative to the other silicate mantle minerals. This increases the silica content of the melt and triggers immediate exsolution of $CO_2$, reducing the buoyancy of the magma and facilitating rapid ascent. The faster the magma rises, the more mantle material will be entrained and the more opx will be dissolved. Providing opx is available, this mechanism enables the continuous and accelerating ascent of the magma and the evolution of the melt from carbonatitic to kimberlitic compositions.

## METHODS SUMMARY

Our experiments use synthetic melts of $Na_2CO_3$ (melting temperature $T_m = 851$ °C) as an analogue for a natural carbonatitic-like melt. The experiments involve the melting of mixtures of $Na_2CO_3$ and crushed natural opx at temperatures of 1,000–1,100 °C for ~1 h; the mixture is open to the air. Pure $Na_2CO_3$ melt contains ~44 wt% dissolved $CO_2$ and is stable at these conditions. Melting of $Na_2CO_3$-opx mixtures produces vigorous decarbonation expressed by weight loss. Decarbonation results from reduced $CO_2$ solubility in the hybrid melt produced by opx assimilation (Fig. 2). Both transient and static experiments were performed. In the former, an initial mass of sample ($Na_2CO_3$ + opx) of known composition was lowered from a balance into a preheated furnace; cumulative weight loss was recorded as a function of time. The transient experiment on pure $Na_2CO_3$ (>1 h at 1,100 °C) shows only slight (<1.8%) change in mass and demonstrates the stability of stoichiometric amounts of $CO_2$ in the melt. Transient melting of $Na_2CO_3$-opx mixtures causes immediate reaction as opx is assimilated; the resulting hybrid melts exsolve $CO_2$ causing a continuous but decreasing loss in weight, until the equilibrium solubility limit of the new melt composition is reached, and weight loss ceases (<20 min). The magnitude of weight loss is taken as a quantitative measure of $CO_2$ loss. The tangents to the weight loss curves give the relative rates of assimilation and degassing. Static experiments saw known masses and compositions of $Na_2CO_3$ + opx placed into a pre-heated oven (1,000–1,100 °C) for a fixed time (Supplementary Table 1; Fig. 3a). The weight loss at the end of each experiment was measured to provide a quantitative estimate of decarbonation.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 29 July; accepted 28 November 2011.**

1. Mitchell, R. H. *Kimberlites: Mineralogy, Geochemistry and Petrology* (Plenum, 1986).
2. Nixon, P. H. The morphology and nature of primary diamondiferous occurrences. *J. Geochem. Explor.* **53,** 41–71 (1995).
3. Kjarsgaard, B. A. in *Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, The Evolution Of Geological Provinces, and Exploration Methods* (ed. Goodfellow, W. D.) 245–271 (Geological Association of Canada, Mineral Deposits Division, 2007).
4. Boyd, F. R. & Nixon, P. H. in *Lesotho Kimberlites* (ed. Nixon, P. H.) 254–268 (Lesotho National Development Corporation, Maseru, 1973).
5. Mitchell, R. H. Petrology of hypabyssal kimberlites: relevance to primary magma compositions. *J. Volcanol. Geotherm. Res.* **174,** 1–8 (2008).
6. Anderson, O. L. in *Kimberlites, Diatremes and Diamonds: Their Geology, Petrology and Geochemistry* (eds Meyer, H. O. A. & Boyd, F. R.) 344–353 (AGU, 1979).
7. Canil, D. & Fedortchouk, Y. Garnet dissolution and the emplacement of kimberlites. *Earth Planet. Sci. Lett.* **167,** 227–237 (1999).
8. Sparks, R. S. J. *et al.* Dynamical constraints on kimberlite volcanism. *J. Volcanol. Geotherm. Res.* **155,** 18 (2006).
9. Wilson, L. & Head, J. W. An integrated model of kimberlite ascent and eruption. *Nature* **447,** 53–57 (2007).
10. Sparks, R. S. J., Pinkerton, H. & Macdonald, R. The transport of xenoliths in magmas. *Earth Planet. Sci. Lett.* **35,** 234–238 (1977).
11. Spera, F. K. Carbon dioxide in petrogenesis III: role of volatiles in the ascent of alkaline magma with special reference to xenolith-bearing mafic lavas. *Contrib. Mineral. Petrol.* **88,** 217–232 (1984).
12. Sparks, R. S. J. *et al.* The nature of erupting kimberlite melts. *Lithos* **112,** 429–438 (2009).
13. Price, S. E., Russell, J. K. & Kopylova, M. G. Primitive magma from the Jericho Pipe, N.W.T., Canada: constraints on primary kimberlite melt chemistry. *J. Petrol.* **41,** 789–808 (2000).
14. Boyd, F. R. Compositional distinction between oceanic and cratonic lithosphere. *Earth Planet. Sci. Lett.* **96,** 15–26 (1989).
15. Kopylova, M. G. & Russell, J. K. Chemical stratification of cratonic lithosphere: constraints from the Northern Slave craton, Canada. *Earth Planet. Sci. Lett.* **181,** 71–87 (2000).
16. McDonough, W. F. Constraints on the composition of the continental lithospheric mantle. *Earth Planet. Sci. Lett.* **101,** 1–18 (1990).
17. Mitchell, R. H. Composition of olivine, silica activity and oxygen fugacity in kimberlite. *Lithos* **6,** 65–81 (1973).
18. Brett, R. C., Russell, J. K. & Moss, S. Origin of olivine in kimberlite: phenocryst or imposter? *Lithos* **112,** 201–212 (2009).
19. Brooker, R. A., Sparks, R. S. J., Kavanagh, J. & Field, M. The volatile content of hypabyssal kimberlite magmas: some constraints from experiments on natural rock compositions. *Bull. Volcanol.* **73,** 959–981 (2011).
20. Sparks, R. S. J., Brown, R. J., Field, M. & Gilbertson, M. A. Kimberlite ascent and eruption. *Nature* **450,** E21 (2007).
21. Dalton, J. A. & Presnall, D. C. The continuum of primary carbonatitic-kimberlitic melt compositions in equilibrium with lherzolite: data from the system CaO-MgO-$Al_2O_3$-$SiO_2$-$CO_2$ at 6 GPa. *J. Petrol.* **39,** 1953–1964 (1998).

22. Dobson, D. P. *et al.* In-situ measurement of viscosity and density of carbonate melts at high pressure. *Earth Planet. Sci. Lett.* **143,** 207–215 (1996).
23. Keppler, H. Water solubility in carbonatite melt. *Am. Mineral.* **88,** 1822–1824 (2003).
24. Lensky, N. G., Niebo, R. W., Holloway, J. R., Lyakhovsky, V. & Navon, O. Bubble nucleation as a trigger for xenolith entrapment in mantle melts. *Earth Planet. Sci. Lett.* **245,** 278–288 (2006).
25. Luth, R. W. The activity of silica in kimberlites: revisited. *Contrib. Mineral. Petrol.* **158,** 283–294 (2009).
26. Shaw, C. S. J. Dissolution of orthopyroxene in basanitic magma between 0.4 and 2 GPa: further implications for the origin of Si-rich alkaline glass inclusions in mantle xenoliths. *Contrib. Mineral. Petrol.* **135,** 114–132 (1999).
27. Edwards, B. R. & Russell, J. K. Time scales of magmatic processes: new insights from dynamic models for magmatic assimilation. *Geology* **26,** 1103–1106 (1998).
28. Wyllie, P. J. & Huang, W. L. Peridotite, kimberlite, and carbonatite explained in the system CaO-MgO-$SiO_2$-$CO_2$. *Geology* **3,** 621–624 (1975).
29. Canil, D. & Bellis, A. J. Phase equilibria in a volatile-free kimberlite at 0.1 MPa and the search for primary kimberlite magma. *Lithos* **105,** 111–117 (2008).
30. Patterson, M., Francis, D. & McCandless, T. Kimberlites: magmas of mixtures? *Lithos* **112,** 191–200 (2009).
31. Kavanagh, J. & Sparks, R. S. J. Temperature changes in ascending kimberlite magma. *Earth Planet. Sci. Lett.* **286,** 404–413 (2009).
32. Dawson, J. B. Quaternary kimberlitic volcanism on the Tanzania Craton. *Contrib. Mineral. Petrol.* **116,** 473–485 (1994).
33. Ferguson, J., Danchin, R. V. & Nixon, P. H. in *Lesotho Kimberlites* (ed. Nixon, P. H.) 207–213 (Lesotho National Development Corporation, Maseru, Lesotho, 1973).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We acknowledge the laboratory support of W. Ertel-Ingritsch and S. Laumann at Ludwig Maximilian University. Funding for this research is from the Natural Sciences and Engineering Research Council (J.K.R.), a Marie Curie outbound fellowship (L.A.P.) and an ERC Advanced Researcher Grant (D.B.D.). The original manuscript benefitted from review by S. Sparks.

**Author Contributions** J.K.R. conceptualized the original idea, performed the static experiments at the University of Munich, and wrote the original draft paper. L.A.P. provided kimberlite expertise and context and drafted the figures. Y.L. performed the transient experiments and processed the experimental data. D.B.D. hosted, provided and repaired the experimental facility, and advised on the optimal experimental methods. All co-authors contributed to producing a final draft for review and in revising the manuscript after review.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.K.R. (krussell@eos.ubc.ca).

doi:10.1038/nature10723

# Recovery rates reflect distance to a tipping point in a living system

Annelies J. Veraart[1], Elisabeth J. Faassen[1], Vasilis Dakos[1], Egbert H. van Nes[1], Miquel Lürling[1,2] & Marten Scheffer[1]

**Tipping points, at which complex systems can shift abruptly from one state to another, are notoriously difficult to predict[1]. Theory proposes that early warning signals may be based on the phenomenon that recovery rates from small perturbations should tend to zero when approaching a tipping point[2,3]; however, evidence that this happens in living systems is lacking. Here we test such 'critical slowing down' using a microcosm in which photo-inhibition drives a cyanobacterial population to a classical tipping point when a critical light level is exceeded. We show that over a large range of conditions, recovery from small perturbations becomes slower as the system comes closer to the critical point. In addition, autocorrelation in the subtle fluctuations of the system's state rose towards the tipping point, supporting the idea that this metric can be used as an indirect indicator of slowing down[4,5]. Although stochasticity prohibits prediction of the timing of critical transitions, our results suggest that indicators of slowing down may be used to rank complex systems on a broad scale from resilient to fragile.**

Systems ranging from the brain and society to ecosystems and the climate can have tipping points at which minor perturbations can invoke a critical transition to a contrasting state[6]. The complexity of such systems prohibits accurate predictive modelling. However, it has been suggested that even without mechanistic insight, the proximity of a tipping point may be inferred from generic features of fluctuations and spatial patterns that can be interpreted as early warning indicators[1,7–9]. This idea is based on the phenomenon that at bifurcation points at which stability of an equilibrium changes, the dominant real eigenvalue becomes zero[10]. As a result, the rate of recovery from perturbation should go to zero as such bifurcations are approached (Supplementary Notes 1). This phenomenon, which is known as critical slowing down, is well established in physics but it has only recently been suggested that the recovery rate from perturbations could be an indicator of the distance to a tipping point in complex living systems such as ecosystems[3].

Although the prospect that the fragility of living systems could be probed this way is attractive, experimental evidence has so far been lacking. Instead, much work has been focused on ways to infer slowing down from indirect indicators such as autocorrelation and variance. However, although these indicators are linked to slowing down in simple stochastically forced models[1,5,9], recent theoretical studies indicate that the indirect indicators will not always respond in simple ways[11] (Brock, W. A. & Carpenter, S. R., submitted). This is confirmed by empirical studies on the climate[5], the food web of a lake[12] and laboratory populations of water fleas[13]. In these systems, trends in indirect indicators occurred but were not all consistent. Here we use a controlled system in which there is positive feedback between organisms and their physical environment to test critical slowing down directly from recovery rates.

We exposed cyanobacteria in chemostat microcosms to increasing light stress. This is a well understood system for which models have shown alternative stable states and tipping points[14,15]. Cyanobacteria provide the shade needed for their own growth, creating a positive feedback, and this constitutes the mechanism behind the bistability. Although light is needed for photosynthesis, light levels that are too high are detrimental to primary producers such as the cyanobacteria that we used. Mutual shading can ameliorate this stress, and is thus one of the ways in which facilitation can outweigh competitive interactions under harsh conditions[16]. Such feedback between organisms and their environment is the mechanism behind alternative stable states in a range of ecosystems[17]. Indeed, as a result of the facilitative shading, our system can maintain a high biomass under incident light levels that prohibit growth in low biomass cultures. A fold bifurcation that represents a classical tipping point occurs at the light level at which this mechanism becomes too weak to allow persistence of the population[14,15] (see Supplementary Notes 1 for a model analysis). Here, shading becomes insufficient to prevent growth inhibition, and the resulting loss of biomass further weakens the stabilizing shading effect. This implies that a positive feedback is driving the system towards a crash.

We cultured cyanobacteria in two independently controlled chemostat microcosms (M1 and M2) and increased incident light daily in small steps to the point at which the population collapsed (see Methods). We perturbed the populations every 4–5 days by removing 10% of their biomass through dilution. Consistent with the model results (Supplementary Fig. 1.1) the populations maintained a relatively high biomass throughout the experiment until they collapsed rather markedly when a tipping point was reached (Fig. 1 and Supplementary Notes 3). Recovery rates of both systems decreased gradually

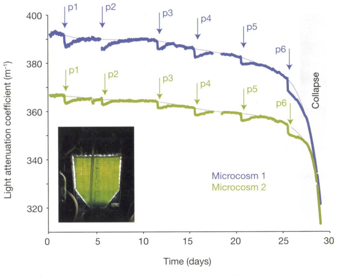


**Figure 1 | The response of two populations of cyanobacteria (*Aphanizomenon flos-aquae*) to dilution events under a regime of gradually increasing light levels.** Dilution events are indicated as perturbations p1–p6. The light attenuation coefficient is a measure of population density. Thin curve segments represent the baselines that were used for computing recovery rates. The inset shows the experimental system.

---

[1]Department of Aquatic Ecology and Water Quality Management, Wageningen University, PO Box 47, NL-6700 AA, Wageningen, The Netherlands. [2]Department of Aquatic Ecology, Netherlands Institute of Ecology, Royal Netherlands Academy of Arts and Sciences, PO Box 50, 6700AB, Wageningen, The Netherlands.

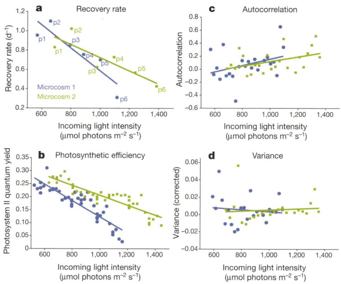


**Figure 2 | Indicators of slowing down as a function of light intensity.** **a**, Recovery rates after perturbation (p1–p6). **b**, Photosynthetic efficiency (photosystem II quantum yield). **c**, Autocorrelation in the population density estimator for each day based on 30 min average $I_{out}$ data. **d**, Corrected variance in daily time series (see Methods and Supplementary Notes 4). Fisher's combined (one-tailed) significance test of the slopes of the regression line's ($n = 6$) in both microcosms: recovery rate, $\chi^2_4 = 18.48$, $P = 0.001$ (if the p6 perturbations are excluded ($n = 5$), $\chi^2_4 = 12.39$, $P = 0.015$); photosynthetic efficiency $n = 45$, $\chi^2_4 = 147.66$, $P < 0.001$; autocorrelation, $n = 23$, $\chi^2_4 = 14.00$, $P = 0.007$; variance, not significant, $n = 23$, $\chi^2_4 = 7.72$, $P = 0.102$.

towards the tipping point, starting far from the bifurcation (Fig. 2a), and tended to decline more rapidly towards the tipping point. This was also predicted by our model (Supplementary Fig. 1.1).

In most complex systems, the mechanisms that are involved in causing the slowing down will be difficult to unravel. However, in our particular system the photosynthetic capacity of the cyanobacteria that is at the heart of their growth potential can be sensed through measurements of the efficiency of their light harvesting system (see Methods). Whereas biomass remained relatively high in the trajectory towards the critical threshold, this specific indicator of their vigour declined linearly with increasing light stress, approaching zero at the point of collapse (Fig. 2b). This is an independent confirmation that the light-induced stress to the cyanobacteria in the two independently operated microcosms does indeed undermine the resilience of the system to the point at which collapse is inevitable.

In systems that are subject to stochastic perturbation regimes, slowing down is predicted to be reflected in characteristic changes in fluctuations of the state. In particular, it has been proposed that increases in autocorrelation and variance can be interpreted as indirect indicators of slowing down[1,4,9]. Although our experiment was not primarily designed to study the effect of stochastic perturbation regimes, there are continuous subtle fluctuations in our measurements of the density of the cyanobacterial population. These fluctuations will reflect a mixture of factors including measurement noise and instabilities in the lighting system as well as true population fluctuations induced by the slightly fluctuating conditions in the bubbled microcosms. We studied how autocorrelation and variance in these subtle fluctuations changed as our system approached the critical point. For this we analysed all the stretches of continuous measurements of 1 day between the experimental interruptions caused by the daily stepwise increase of light intensity and the dilution perturbation events (see Methods). Although autocorrelation was quite variable in the time series that we studied, there was a significant increase towards the tipping point in both experimental systems (Fig. 2c). No trend in variance was apparent in the systems (Fig. 2d, see also Supplementary Notes 4). These results are consistent with the prediction that measuring recovery rates from perturbations is a robust way to detect critical slowing down[18], whereas indirect indicators of slowing down may or may not increase towards a critical threshold[11]. Also, our findings are consistent with the prediction that autocorrelation is usually directly related to slowing down and may therefore provide a more robust signal than variance in some situations[11].

Perturbation experiments such as the ones in our experiment will often be impossible in large complex systems, leaving indirect indicators as the only tool by which to infer slowing down. However, experimental probing of recovery rates may be feasible in some smaller systems, as long as the timescales are appropriate and stochastic fluctuations are small relative to experimental perturbations. Even in larger systems, local perturbations may be an option to probe resilience, allowing adaptive management to steer the system away from the brink of collapse.

Perhaps most importantly, our experimental demonstration of slowing down implies a proof of concept, providing a fundamental basis to the current search for generic early warning signals in systems ranging from the brain and ecosystems to society and the climate[1]. The fact that slowing down in our system started far from the critical point suggests that recovery rates as well as indirect indicators may be used to rank such complex systems on a broad scale from stable to critical. This does not mean that slowing down can be used to actually predict transitions. Stochastic shocks will trigger critical transitions always before the bifurcation point is reached, indicating that there is inherent unpredictability in systems. Nevertheless, the prospect of having generic indicators of resilience is a potentially large step forward. Mechanistic models to predict tipping points in nature and society accurately are simply beyond our reach, leaving empirical estimation of fragility as one of the key challenges in complex systems science today[1].

## METHODS SUMMARY

Experiments were performed in two identical flat chemostat microcosms[19] (M1 and M2) in which we cultured cyanobacteria (*Aphanizomenon flos-aquae* (L.) Ralfs) on a nutrient-rich growth medium, modified from BG11 medium[20]. Light irradiance was increased in steps of 23 μmol photons m$^{-2}$ s$^{-1}$ for M1 and 29 μmol photons m$^{-2}$ s$^{-1}$ each day for M2. Photosynthetic efficiency was measured from diurnal samples. The intensity of the light passing through the chemostats was averaged at 5-min intervals, and light attenuation coefficients (ε, m$^{-1}$) were calculated as an indicator of biomass:

$$\varepsilon = \frac{-\ln(I_{out}/I_{in})}{d}$$

where $I_{in}$ is the intensity of the incoming light (μE), $I_{out}$ is the intensity of the outgoing light (μE) and $d$ is the depth of the chemostats (m). External perturbations were performed every 4–5 days by diluting the culture with 170 ml of sterile medium. A baseline for calculating recovery rates was constructed for each perturbation event (Supplementary Notes 2 and Supplementary Fig. 2.1) by fitting a quadratic curve from the period just before perturbation to the period just before the next perturbation (thin curves in Fig. 1). Recovery rates after each perturbation (λ, *per day*) were then calculated from a linear regression of $-\ln(\varepsilon_0 - \varepsilon_c)$ against time, where $\varepsilon_0$ is the light attenuation of the baseline (m$^{-1}$) and $\varepsilon_c$ is the light attenuation of the chemostat (m$^{-1}$)[3].

The lag 1 autocorrelation and variance were analysed for each uninterrupted period between the daily manipulations after removing the trends from each period by fitting polynomials of 2 degrees. To check for effects of nonlinear propagation of measurement noise, we constructed a null model that assumed that all residuals were due to uncorrelated normally distributed noise. As variance showed a trend towards the bifurcation in this null model (Supplementary Table 4.1), we corrected the observed variance by subtracting the median of the null model (Fig. 2d).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 13 September; accepted 18 November 2011.**

**Published online 25 December 2011.**

1. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461,** 53–59 (2009).

2. Wissel, C. A universal law of the characteristic return time near thresholds. *Oecologia* **65,** 101–107 (1984).
3. van Nes, E. H. & Scheffer, M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am. Nat.* **169,** 738–747 (2007).
4. Ives, A. R. Measuring resilience in stochastic systems. *Ecol. Monogr.* **65,** 217–233 (1995).
5. Dakos, V. *et al.* Slowing down as an early warning signal for abrupt climate change. *Proc. Natl Acad. Sci. USA* **105,** 14308–14312 (2008).
6. Scheffer, M. *Critical Transitions in Nature and Society* (eds Levin, S.A. & Strogatz, S.H.) (Princeton Univ. Press, 2009).
7. Kleinen, T., Held, H. & Petschel-Held, G. The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. *Ocean Dyn.* **53,** 53–63 (2003).
8. Held, H. & Kleinen, T. Detection of climate system bifurcations by degenerate fingerprinting. *Geophys. Res. Lett.* **31,** L23207 (2004).
9. Carpenter, S. R. & Brock, W. A. Rising variance: a leading indicator of ecological transition. *Ecol. Lett.* **9,** 311–318 (2006).
10. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering* 1st edn (Addison-Wesley, 1994).
11. Dakos, V., van Nes, E. H., D' Odorico, P. & Scheffer, M. How robust are variance and autocorrelation as early-warning signals for critical transitions? *Ecology* (submitted); preprint at http://dx.doi.org/10.1890/11-0889.1.
12. Carpenter, S. R. *et al.* Early warnings of regime shifts: A whole-ecosystem experiment. *Science* **332,** 1079–1082 (2011).
13. Drake, J. M. & Griffen, B. D. Early warning signals of extinction in deteriorating environments. *Nature* (2010).
14. Huisman, J. *The Struggle for Light.* PhD thesis, Univ. Groningen (1997).
15. Gerla, D. J., Mooij, W. M. & Huisman, J. Photoinhibition and the assembly of light-limited phytoplankton communities. *Oikos* **120,** 359–368 (2011).
16. Holmgren, M., Scheffer, M. & Huston, M. A. The interplay of facilitation and competition in plant communities. *Ecology* **78,** 1966–1975 (1997).
17. Scheffer, M., Carpenter, S. R., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413,** 591–596 (2001).
18. Dakos, V., Kéfi, S., Rietkerk, M., van Nes, E. H. & Scheffer, M. Slowing down in spatially patterned ecosystems at the brink of collapse. *Am. Nat.* **177,** E153–E166 (2011).
19. Huisman, J. *et al.* Principles of the light-limited chemostat: theory and ecological applications. Antonie Leeuwenhoek **81,** 117–133 (2002).
20. Andersen, R. A. Berges, J.A., Harrison, P.J. & Watanabe M.M. in *Algal culturing techniques* 1st edn 435–436 (Elsevier, 2005).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank M. B. Gonçalves Souza for discussions on the experimental set up. We thank D. Waasdorp and W. Beekman-Lukassen for assistance with the experiments and C. ter Braak for statistical advice. A.J.V., E.J.F., V.D., E.H.v.N. and M.S. are supported by a European Research Council Advanced grant and M.S. is the recipient of a Spinoza award.

**Author Contributions** A.J.V., E.J.F. and M.L. performed the experiments. A.J.V., E.J.F, E.H.v.N. and V.D. analysed the data. M.S., A.J.V., E.H.v.N., E.J.F. and V.D. wrote the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.H.v.N. (egbert.vannes@wur.nl).

doi:10.1038/nature10724

# Evolution of increased complexity in a molecular machine

Gregory C. Finnigan[1]*, Victor Hanson-Smith[2,3]*, Tom H. Stevens[1] & Joseph W. Thornton[2,4,5]

**Many cellular processes are carried out by molecular 'machines'—assemblies of multiple differentiated proteins that physically interact to execute biological functions[1–8]. Despite much speculation, strong evidence of the mechanisms by which these assemblies evolved is lacking. Here we use ancestral gene resurrection[9–11] and manipulative genetic experiments to determine how the complexity of an essential molecular machine—the hexameric transmembrane ring of the eukaryotic V-ATPase proton pump—increased hundreds of millions of years ago. We show that the ring of Fungi, which is composed of three paralogous proteins, evolved from a more ancient two-paralogue complex because of a gene duplication that was followed by loss in each daughter copy of specific interfaces by which it interacts with other ring proteins. These losses were complementary, so both copies became obligate components with restricted spatial roles in the complex. Reintroducing a single historical mutation from each paralogue lineage into the resurrected ancestral proteins is sufficient to recapitulate their asymmetric degeneration and trigger the requirement for the more elaborate three-component ring. Our experiments show that increased complexity in an essential molecular machine evolved because of simple, high-probability evolutionary processes, without the apparent evolution of novel functions. They point to a plausible mechanism for the evolution of complexity in other multi-paralogue protein complexes.**

Comparative genomic approaches suggest that the components of many molecular machines have appeared sequentially during evolution and that complexity increased gradually by incorporating new parts into simpler assemblies[2–8]. Such horizontal analyses of extant systems, however, cannot decisively test these hypotheses or reveal the mechanisms by which additional parts became obligate components of larger complexes. In contrast, vertical approaches that combine computational phylogenetic analysis with gene synthesis and molecular assays allow changes in the sequence, structure and function of reconstructed ancestral proteins to be experimentally traced through time.[9–11] Here we apply this approach to characterize the evolution of a small molecular machine and dissect the mechanisms that caused it to increase in complexity.

The vacuolar H$^+$-ATPase (V-ATPase) is a multisubunit protein complex that pumps protons across membranes to acidify subcellular compartments; this function is required for intracellular protein trafficking, coupled transport of small molecules and receptor-mediated endocytosis[1]. V-ATPase dysfunction has been implicated in human osteoporosis, in acquired drug resistance in human tumours, and in pathogen virulence[12–14]. A key subcomplex of the V-ATPase is the V$_0$ protein ring, a hexameric assembly that uses a rotary mechanism to move protons across organelle membranes (Fig. 1a)[15,16]. Although the V-ATPase is found in all eukaryotes, the V$_0$ ring varies in subunit composition among lineages. In animals and most other eukaryotes,

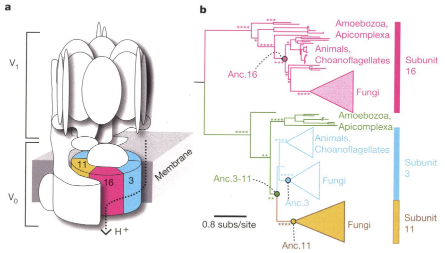


**Figure 1 | Structure and evolution of the V-ATPase complex. a**, In *S. cerevisiae*, the V-ATPase contains two subcomplexes: the octameric V$_1$ domain is on the cytosolic side of the organelle membrane, and the hexameric V$_0$ ring is membrane bound. Protein subunits Vma3, Vma11 and Vma16 are labelled and coloured. **b**, Maximum likelihood phylogeny of V-ATPase subunits Vma3, Vma11 and Vma16. All eukaryotes contain subunits 3 and 16, but Fungi also contain subunit 11. Circles show ancestral proteins reconstructed in this study. Colours correspond to those of subunits in panel **a**; unduplicated orthologues of Vma3 and Vma11 are green. Asterisks show approximate likelihood ratios for major nodes: ****, $>10^3$; ***, $>10^2$; **, $>10$; *, $<10$; ~, $<2$. The complete phylogeny is presented in Supplementary Information, section 2.

---

[1] Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA. [2] Institute for Ecology and Evolution, University of Oregon, Eugene, Oregon 97403, USA. [3] Department of Computer and Information Science, University of Oregon, Eugene, Oregon 97403, USA. [4] Howard Hughes Medical Institute, Eugene, Oregon 97403, USA. [5] Departments of Human Genetics and Ecology & Evolution, University of Chicago, Chicago, Illinois 60637, USA.
*These authors contributed equally to this work.

the ring consists of one subunit of Vma16 protein and five copies of its paralogue, Vma3 (Fig. 1b)[1]. In Fungi, the ring consists of one Vma16 subunit, four copies of Vma3 and one Vma11 subunit, arranged in a specific orientation[17]. All three proteins are required for V-ATPase function in Fungi[18,19], but the mechanisms are unknown by which both Vma3 and Vma11 became obligate components with specific positional roles in the complex.

To understand how the three-component ring evolved, we reconstructed ancestral $V_0$ proteins from just before and after the increase in complexity, synthesized and functionally characterized them in a yeast genetic system, and used manipulative methods to identify the genetic and molecular mechanisms by which their functions changed. We first inferred the phylogeny and best-fit evolutionary model of the protein family of which Vma3, Vma11 and Vma16 are members, using the sequences of all 139 extant family members available in GenBank (Supplementary Table 1). The maximum likelihood phylogeny (Fig. 1b and Supplementary Information, section 2) indicates that Vma3 and Vma11 are sister proteins that were produced by duplication of an ancestral gene (Anc.3-11) before the last common ancestor of all Fungi (~800 million years ago[20]). Whether this duplication occurred before or after the divergence of Fungi from other eukaryotes (~1 billion years ago[20]) is not clearly resolved, although the latter scenario is more parsimonious. The Vma3/Vma11 and Vma16 lineages, in turn, descend from an older gene duplication deep in the eukaryotic lineage (Fig. 1b). We used a maximum likelihood algorithm[21] to infer the ancestral amino acid sequences with the highest probability of producing all the extant sequence data, given the best-fit phylogeny and model. We reconstructed the ancestral proteins (Anc.3-11 and Anc.16) that made up the ancient two-paralogue eukaryotic ring, as well as the duplicated subunits Anc.3 and Anc.11 from the three-component ring in the common ancestor of all Fungi (Supplementary Information, sections 3 and 4).

To characterize the functions of these reconstructed proteins, we synthesized coding sequences and transformed them into *Saccharomyces cerevisiae* deficient for various ring components and therefore incapable of growth in the presence of elevated $CaCl_2$ (ref. 22). We found that the ancestral two-subunit ring can functionally replace the three-subunit ring of extant yeast. When the resurrected Anc.3-11 was transformed into yeast deficient for Vma3 (*vma3Δ*) or Vma11(*vma11Δ*), growth in the presence of elevated $CaCl_2$ was rescued, indicating that the functions of the present-day Vma3 and Vma11 proteins were already present before the duplication that generated them (Fig. 2a). Furthermore, Anc.3-11—unlike either of its present-day descendants—can partially rescue growth in yeast that are doubly deficient for both Vma3 and Vma11 (*vma3Δ vma11Δ*). The reconstructed Anc.16 also rescued growth in Vma16-deficient *S. cerevisiae* (*vma16Δ*) (Fig. 2b), and co-expression of Anc.3-11 and Anc.16 together rescued cell growth in *vma3Δ vma11Δ vma16Δ* yeast, which lack all three ring subunits (Fig. 2c). The ancestral genes specifically restore proper V-ATPase function in acidification of the vacuolar lumen (Fig. 2g). In addition, mutation of the ancestral subunits to remove glutamic acid residues known to be essential for V-ATPase enzyme function[17,23] abolished their ability to rescue growth on $CaCl_2$ (Supplementary Information, section 7). These inferences about the functions of Anc.3-11 and Anc.16 are robust to uncertainty about ancestral amino acid states. We reconstructed alternative versions of Anc.3-11 and Anc.16 by introducing amino acid states with posterior probability >0.2, but none of these abolished the ability of the ancestral genes to substitute functionally for the extant subunits (Supplementary Information, section 8). These results establish that during the increase in complexity, neither the $V_0$ complex nor its component proteins evolved new functions required for growth under the conditions in which the ring is known to be important.

Similar experiments with the components of the ancestral three-component ring show that after the duplication of Anc.3-11, its descendants Anc.3 and Anc.11 both became necessary for a functional complex because of complementary losses of ancestral functions. Unlike Anc.3-11, expression of Anc.3 can rescue growth and vacuole acidification in *vma3Δ* but not *vma11Δ* yeast, and Anc.11 can rescue growth in *vma11Δ* but not *vma3Δ* yeast (Fig. 2d, e, g). Furthermore, both Anc.3 and Anc.11 are required to rescue growth fully in *vma3Δ vma11Δ* yeast (Fig. 2f). These data indicate that after its origin by gene duplication, Anc.11 lost the ancestral protein's ability to carry out at least some functions of Vma3, and Anc.3 lost the ancestral capacity to carry out those of Vma11.

We conjectured that Vma3 and Vma11 evolved their specialized roles because they lost specific interfaces present in their ancestor that are required for ring assembly. Previous experiments with fusions of

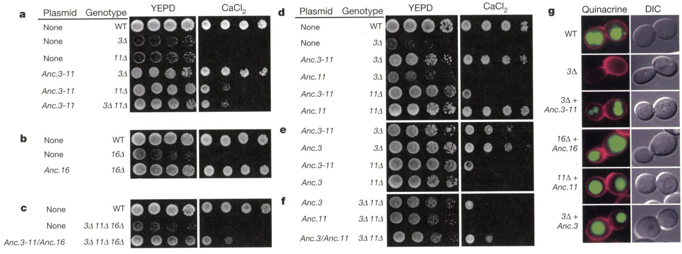


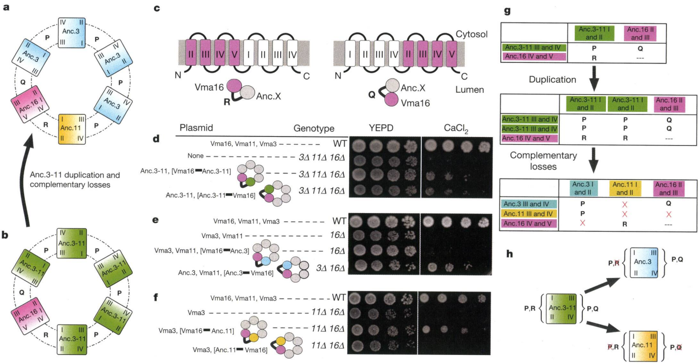
**Figure 2 | Two reconstructed ancestral $V_0$ subunits functionally replace the three-paralogue ring in extant yeast.** *S. cerevisiae* were plated in decreasing concentrations on permissive medium (YEPD) buffered with elevated $CaCl_2$. **a**, Expression of Anc.3-11 rescues growth in yeast that are deficient for endogenous subunit Vma3 (*3Δ*), subunit Vma11 (*11Δ*) or both (*3Δ 11Δ*). Growth of wild-type (WT) yeast is shown for comparison. **b**, Anc.16 rescues growth in yeast that are deficient for subunit Vma16 (*16Δ*). **c**, Expression of Anc.3-11 and Anc.16 together rescues growth in yeast that are deficient for Vma3, Vma11 and Vma16. **d**, Anc.11 rescues growth in *vma11Δ* but not in *vma3Δ* yeast. **e**, Anc.3 rescues growth in *vma3Δ* but not *vma11Δ* yeast. **f**, Anc.3 and Anc.11 together rescue growth in *vma3Δ vma11Δ* mutants. **g**, Yeast expressing reconstructed ancestral subunits properly acidified the vacuolar lumen. Red signal shows yeast cell walls; green signal (quinacrine) shows acidified compartments. Yeast were visualized by differential interference contrast microscopy.

extant yeast proteins have shown that the arrangement of subunits in the ring is constrained by the capacity of each subunit to form specific interfaces (which we labelled P, Q and R) with the other subunits[24]. Specifically, Vma11 is restricted to a single position between Vma16 and Vma3, because its clockwise interface can participate only in interface R with Vma16, and its anticlockswise interface can participate only in interface P with the clockwise side of Vma3 (Fig. 3). By contrast, copies of Vma3 occupy several positions in the ring, because they form interface P with other copies of Vma3 or Vma11, as well as interface Q with Vma16. However, Vma3 cannot form interface R with Vma16. As a result, both Vma3 and Vma11 are required in extant yeast to form a complete ring with Vma16.

To determine whether interaction interfaces were lost during evolution, we engineered fusions of ancestral ring proteins to assess the capacity of each to form the specific interfaces with other subunits that are required for a functional complex. Because Anc.3-11 can complement the loss of both Vma3 and Vma11, we proposed that the Anc.3-11 subunit could participate in all three specific interaction interfaces, and that these capacities were then partitioned between Anc3 and Anc11 after the duplication of Anc.3-11 (Fig. 3a, b). To test this hypothesis, we created six reciprocal gene fusions between yeast subunit Vma16 and ancestral subunits Anc.3-11, Anc.3 and Anc.11 (Fig. 3c and Supplementary Information, section 9). Each fusion constrains the structural position of subunits relative to subunit Vma16, making it possible to determine which arrangements yield a functional ring. As predicted, Anc.3-11 functioned on either side of Vma16 (Fig. 3d), indicating that it could form all three interfaces P, Q and R. By contrast, Anc.3 functioned when constrained to participate in interface Q with Vma16 and interface P with Vma3; however, ring function was lost when Anc.3 was constrained to form interface R with Vma16 (Fig. 3e). Anc.11 functioned when constrained to participate in interface R with Vma16 and interface P with Vma3, but ring function was lost when Anc.11 was constrained to participate in interface Q with Vma16 and interface P with Vma3. This result indicates that Anc.11 lost the capacity to form one or both of these interfaces during its post-duplication divergence from Anc.3-11 (Fig. 3f).

Taken together, these data indicate that the specificity of the ring arrangement and the obligate roles of Vma3 and Vma11 evolved by complementary loss of asymmetric interactions with other members of the ring (Fig. 3g, h). Before Anc.3-11 duplicated, the protein ring contained copies of only undifferentiated subunit Vma3/Vma11 and subunit 16. Immediately after Anc.3-11 duplicated, the two descendant subunits must have been functionally identical, so the protein ring could have assembled with many possible combinations of the two descendants, including copies of only one of the descendant proteins. This flexibility disappeared when Anc.3 lost the ancestral interface that allowed it to interact with the anticlockwise side of Vma16, and Anc.11 lost the ability to interact with the clockwise side of Vma16 and/or the anticlockwise side of Vma3. These complementary losses are sufficient to explain the specific arrangement of contemporary subunits in reconstructed and present-day fungal transmembrane rings.

To establish the genetic basis for the partitioning of the functions of Anc.3-11 between Vma3 and Vma11, we introduced historical mutations into Anc3.11 by directed mutagenesis and determined whether they recapitulated the shifts in function that occurred during the evolution of Anc.3 and Anc.11. The two phylogenetic branches leading from Anc.3-11 to Anc.3 and to Anc.11 contain 25 and 31 amino acid substitutions, respectively, but only a subset of these are strongly conserved in subunits Vma3 or Vma11 from extant Fungi (Fig. 4a). We introduced each of these 'diagnostic' substitutions into Anc.3-11 and

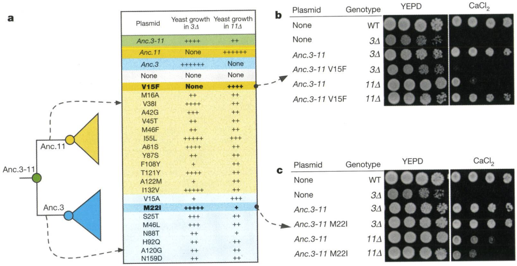


| | Anc.3-11 I and II | Anc.16 II and III |
|---|---|---|
| Anc.3-11 III and IV | P | Q |
| Anc.16 IV and V | R | --- |

| | Anc.3-11 I and II | Anc.3-11 I and II | Anc.16 II and III |
|---|---|---|---|
| Anc.3-11 III and IV | P | P | Q |
| Anc.3-11 III and IV | P | P | Q |
| Anc.16 IV and V | R | R | --- |

| | Anc.3 I and II | Anc.11 I and II | Anc.16 II and III |
|---|---|---|---|
| Anc.3 III and IV | P | ✗ | Q |
| Anc.11 III and IV | P | ✗ | ✗ |
| Anc.16 IV and V | ✗ | R | --- |

**Figure 3 | Increasing complexity by complementary loss of interactions in the fungal $V_0$ ring.** **a**, Model of the ancestral three-paralogue ring, arranged as in extant yeast[24]. Unique intersubunit interfaces are labelled P, Q and R. Subunits are colour-coded as in Fig. 1. **b**, Model of the ancestral two-paralogue ring, before duplication of Anc.3-11. **c**, To constrain the location of specific subunits, gene fusions were constructed by tethering an ancestral subunit to either the amino- or carboxy-terminal side of yeast Vma16. Roman numerals indicate the locations of transmembrane helices (I, II, III, IV and V)[24]. **d**–**f**, Growth assays of yeast with fused $V_0$ subunits identify the interfaces that ancestral subunits can form. For each experiment, expressed $V_0$ subunits are listed. Tethered subunits are in brackets and connected by a thick line. Cartoons show the constrained location of the tethered subunit relative to Vma16. Anc.3-11 can function on either side of Vma16 (**d**). Anc.3 can function only on the clockwise side of Vma16 (**e**). Anc.11 can function only on the anticlockwise side of Sc.16 (**f**). **g**, Interfaces that are formed by $V_0$ subunits before and after duplication and complementary loss of interfaces, based on the data in panels **d**–**f**. Red crosses indicate lost interfaces. **h**, Schematic of interfaces formed by Anc.3-11 that were lost in Anc.3 and Anc.11, based on data in panels **d**–**f**.



| Plasmid | Yeast growth in *3Δ* | Yeast growth in *11Δ* |
|---|---|---|
| *Anc.3-11* | ++++ | ++ |
| *Anc.11* | None | ++++++ |
| *Anc.3* | ++++++ | None |
| None | None | None |
| **V15F** | **None** | **++++** |
| M16A | ++ | ++ |
| V38I | ++++ | ++ |
| A42G | +++ | ++ |
| V45T | ++ | ++ |
| M46F | ++ | ++ |
| I55L | +++++ | +++ |
| A61S | ++++ | ++ |
| Y87S | ++ | ++ |
| F108Y | + | ++ |
| T121Y | ++++ | ++ |
| A122M | + | ++ |
| I132V | +++++ | ++ |
| V15A | + | +++ |
| **M22I** | **+++++** | **+** |
| S25T | +++ | ++ |
| M46L | +++ | ++ |
| N88T | ++ | + |
| H92Q | ++ | ++ |
| A120G | ++ | ++ |
| N159D | ++ | ++ |

**Figure 4 | Genetic basis for functional differentiation of Anc.3 and Anc.11.** **a**, Experimental analysis of historical amino acid replacements. The table lists replacements that occurred on the branches leading from Anc.3-11 to Anc.11 (yellow) or to Anc.3 (blue) and that were subsequently conserved. Each derived residue was introduced singly into Anc.3-11; the variant genes were transformed into *S. cerevisiae*, and growth was assayed on elevated $CaCl_2$. The table shows growth semiquantitatively from zero (none) to wild type (++++++). Bold mutations entirely or partly recapitulate the functional evolution of Anc.11 and Anc.3. **b**, Replacement V15F abolishes the capacity of Anc.3-11 to function as subunit 3 and enhances the capacity of Anc.3-11 to function as subunit 11. **c**, Replacement M22I impairs the capacity of Anc.3-11 to function as subunit 11 without affecting its capacity to function as subunit 3.

experimentally evaluated whether they recapitulated the loss by Anc.3 or Anc.11 of the capacity to complement Vma gene deletions. We found that a single amino-acid replacement that occurred on the branch leading to Anc.11 (V15F) abolished the capacity of Anc.3-11 to function as subunit 3; it also enhanced the ability of Anc.3-11 to function as subunit 11 (Fig. 4b). V15F is located in transmembrane helix I, which participates in the P interface that our experiments indicate may have been lost on the same branch (Fig. 3 and Supplementary Information, section 4). Conversely, a single historical replacement (M22I) on the branch leading to Anc.3 radically reduced the capacity of Anc.3-11 to function as subunit 11 (Fig. 4c). M22I is also in transmembrane helix I, which participates in formation of the R interface that was lost on this branch (Fig. 3 and Supplementary Information, section 4). The Anc.3-11 M22I mutant retains some of the capacity of the ancestral protein to rescue growth in the Vma11-deficient background, suggesting that other mutations also contributed to the functional evolution of Vma3. One other historical mutation (N88T) on this branch also impaired the capacity of Anc.3-11 to function as subunit 11, but it reduced the capacity of the protein to function as Vma3 as well, suggesting that epistatic interactions with other residues allow this mutation to be tolerated in Anc.3 and its descendants. Several of the replacements on the branch leading to Anc.11 show a similar pattern, reducing the capacity of the protein to replace Vma3, indicating that these historical replacements function better together than in isolation.

How complexity and novel functions evolve has been a longstanding question in evolutionary biology[25–27], because mutations that compromise existing functions are far more frequent than those that generate new ones[28]. Our results indicate that the architectural complexity of molecular assemblies can evolve because of a few simple, relatively high-probability mutations that degrade ancestral interfaces but leave other functions intact. The specific roles of subunits Vma3 and Vma11 seem to have been acquired when duplicated genes lost some, but not all, of the capacity of the ancestral protein to participate in interactions with copies of itself and another protein required for proper ring assembly. Because complementary losses occurred in both lineages, the two descendant subunits became obligate components, and the complexity of the ring increased. It is possible that specialization of the duplicated subunits allowed increases in fitness, but genome-wide interaction screens and the phenotype of *vma11Δ* yeast provide no evidence that Vma11 evolved novel functions in addition to those that it inherited from Anc.3-11 in the $V_0$ ring[29].

We are aware of no other mechanistic analyses of a molecular machine's evolutionary trajectory, so the generality of our observations is unknown. By definition, however, all molecular machines involve differentiated parts in specific spatial orientations, and many such complexes are entirely or partially composed of paralogous proteins[2–8]. In the evolution of any such assembly, additional paralogues could become obligate components because of gene duplication[30] and subsequent mutations that cause specific interaction interfaces among them to degenerate.

This view of the evolution of molecular machines is related to recent models that explain other biological phenomena—such as the retention of large numbers of duplicate genes and mobile genetic elements within genomes—as the product of degenerative processes acting on modular biological systems[27]. Although mutations that enhanced the functions of individual ring components may have occurred during evolution, our data indicate that simple degenerative mutations are sufficient to explain the historical increase in complexity of a crucial molecular machine. There is no need to invoke the acquisition of 'novel' functions caused by low-probability mutational combinations.

## METHODS SUMMARY

Ancestral protein sequences were inferred using maximum-likelihood phylogenetics from an alignment of 139 protein sequences of extant subunits 3, 11 and 16 from Amoebozoa, Apicomplexa, Metazoa, Choanoflagellida and Fungi. Ancestral genes were synthesized, cloned into yeast expression vectors and tested for complementation in various *S. cerevisiae* mutants. V-ATPase function was assayed by growth tests on medium buffered with $CaCl_2$, as described previously[31]. Steady-state levels of Vph1 were determined by western blot. Quinacrine staining and Vph1-GFP (green fluorescent protein) fusion constructs were visualized by fluorescence microscopy.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 21 September; accepted 21 November 2011.**
**Published online 9 January 2012.**

1. Forgac, M. Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. *Nature Rev. Mol. Cell Biol.* **8,** 917–929 (2007).

2. Pallen, M. J. & Matzke, N. J. From the origin of species to the origin of bacterial flagella. *Nature Rev. Microbiol.* **4,** 784–790 (2006).
3. Liu, R. & Ochman, H. Stepwise formation of the bacterial flagellar system. *Proc. Natl Acad. Sci. USA* **104,** 7116–7121 (2007).
4. Mulkidjanian, A. Y., Makarova, K. S., Galperin, M. Y. & Koonin, E. V. Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nature Rev. Microbiol.* **5,** 892–899 (2007).
5. Dolezal, P., Likic, V., Tachezy, J. & Lithgow, T. Evolution of the molecular machines for protein import into mitochondria. *Science* **313,** 314–318 (2006).
6. Clements, A. *et al.* The reducible complexity of a mitochondrial molecular machine. *Proc. Natl Acad. Sci. USA* **106,** 15791–15795 (2009).
7. Archibald, J. M., Logsdon, J. M. Jr & Doolittle, W. F. Origin and evolution of eukaryotic chaperonins: phylogenetic evidence for ancient duplications in CCT genes. *Mol. Biol. Evol.* **17,** 1456–1466 (2000).
8. Gabaldón, T., Rainey, D. & Huynen, M. A. Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (complex I). *J. Mol. Biol.* **348,** 857–870 (2005).
9. Thornton, J. W. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nature Rev. Genet.* **5,** 366–375 (2004).
10. Liberles, D. (ed.) *Ancestral Sequence Reconstruction* (Oxford Univ. Press, 2007).
11. Harms, M. J. & Thornton, J. W. Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.* **20,** 360–366 (2010).
12. Frattini, A. *et al.* Defects in TCIRG1 subunit of the vacuolar proton pump are responsible for a subset of human autosomal recessive osteopetrosis. *Nature Genet.* **25,** 343–346 (2000).
13. Pérez-Sayáns, M., Somoza-Martìn, J. M., Barros-Angueira, F., Rey, J. M. & Garcìa-Garcìa, A. V-ATPase inhibitors and implication in cancer treatment. *Cancer Treat. Rev.* **35,** 707–713 (2009).
14. Xu, L. *et al.* Inhibition of host vacuolar H$^+$-ATPase activity by a *Legionella pneumophila* effector. *PLoS Pathog.* **6,** e1000822 (2010).
15. Hirata, T. *et al.* Subunit rotation of vacuolar-type proton pumping ATPase: relative rotation of the g and c subunits. *J. Biol. Chem.* **278,** 23714–23719 (2003).
16. Imamura, H. *et al.* Rotation scheme of V$_1$-motor is different from that of F$_1$-motor. *Proc. Natl Acad. Sci. USA* **102,** 17929–17933 (2005).
17. Powell, B., Graham, L. A. & Stevens, T. H. Molecular characterization of the yeast vacuolar H$^+$-ATPase proton pore. *J. Biol. Chem.* **275,** 23654–23660 (2000).
18. Umemoto, N., Yoshihisa, T., Hirata, R. & Anraku, Y. Roles of the *VMA3* gene product, subunit c of the vacuolar membrane H$^+$-ATPase on vacuolar acidification and protein transport. A study with *VMA3*-disrupted mutants of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **265,** 18447–18453 (1990).
19. Umemoto, N., Ohya, Y. & Anraku, Y. *VMA11*, a novel gene that encodes a putative proteolipid, is indispensable for expression of yeast vacuolar membrane H$^+$-ATPase activity. *J. Biol. Chem.* **266,** 24526–24532 (1991).
20. Taylor, J. W. & Berbee, M. L. Dating divergences in the fungal tree of life: review and new analyses. *Mycologia* **98,** 838–849 (2006).
21. Yang, Z., Kumar, S. & Nei, M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141,** 1641–1650 (1995).
22. Kane, P. M. The where, when, and how of organelle acidification by the yeast vacuolar H$^+$-ATPase. *Microbiol. Mol. Biol. Rev.* **70,** 177–191 (2006).
23. Hirata, R., Graham, L. A., Takatsuki, A., Stevens, T. H. & Anraku, Y. Vma11 and vma16 encode second and third proteolipid subunits of the *Saccharomyces cerevisiae* vacuolar membrane H$^+$-ATPase. *J. Biol. Chem.* **272,** 4795–4803 (1997).
24. Wang, Y., Cipriano, D. J. & Forgac, M. Arrangement of subunits in the proteolipid ring of the V-ATPase. *J. Biol. Chem.* **282,** 34058–34065 (2007).
25. Ohno, S. *Evolution by Gene Duplication* (Springer, 1970).
26. Jacob, F. Evolution and tinkering. *Science* **196,** 1161–1166 (1977).
27. Lynch, M. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc. Natl Acad. Sci. USA* **104,** 8597–8604 (2007).
28. Hietpas, R. T., Jensen, J. D. & Bolon, D. N. Experimental illumination of a fitness landscape. *Proc. Natl Acad. Sci. USA* **108,** 7896–7901 (2011).
29. Tong, A. H. Y. *et al.* Global mapping of the yeast genetic interaction network. *Science* **303,** 808–813 (2004).
30. Pereira-Leal, J. B., Levy, E. D., Kamp, C. & Teichmann, S. A. Evolution of protein complexes by duplication of homomeric interactions. *Genome Biol.* **8,** R51 (2007).
31. Ryan, M., Graham, L. A. & Stevens, T. H. Voa1p functions in V-ATPase assembly in the yeast endoplasmic reticulum. *Mol. Biol. Cell* **19,** 5131–5142 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** This study was supported by National Institutes of Health (NIH) grants R01-GM081592 (to J.W.T.) and R01-GM38006 (to T.H.S.), National Science Foundation (NSF) grants IOB-0546906 (to J.W.T.) and DEB-0516530 (to J.W.T.), NIH Genetics Training grant T32-GM007257 (to G.C.F.), NSF IGERT grant DGE-9972830 (to V.H.-S.) and the Howard Hughes Medical Institute (J.W.T.). We thank L. Graham, G. Butler and B. Houser for generating yeast strains and other assistance. We thank members of the Stevens and Thornton laboratories for helpful comments.

**Author Contributions** V.H.-S. performed the phylogenetic analysis and statistical reconstructions. G.C.F. performed functional experiments. All authors conceived the experiments, interpreted the results and wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.W.T. (joet@uoregon.edu).

doi:10.1038/nature10719

# Global landscape of HIV–human protein complexes

Stefanie Jäger[1,2], Peter Cimermancic[2,3], Natali Gulbahce[1,2], Jeffrey R. Johnson[1,2,4], Kathryn E. McGovern[1,2], Starlynn C. Clarke[5], Michael Shales[1,2], Gaelle Mercenne[6], Lars Pache[7], Kathy Li[1,2,5], Hilda Hernandez[1,2,5], Gwendolyn M. Jang[1,2,8], Shoshannah L. Roth[9], Eyal Akiva[2,3], John Marlett[10], Melanie Stephens[9], Iván D'Orso[8]†, Jason Fernandes[8], Marie Fahey[1,2], Cathal Mahon[1,2,5], Anthony J. O'Donoghue[5], Aleksandar Todorovic[11], John H. Morris[5], David A. Maltby[5], Tom Alber[12], Gerard Cagney[13], Frederic D. Bushman[9], John A. Young[10], Sumit K. Chanda[7], Wesley I. Sundquist[6], Tanja Kortemme[2,3,14], Ryan D. Hernandez[2,3,14], Charles S. Craik[2,5], Alma Burlingame[2,5], Andrej Sali[2,3,5,14], Alan D. Frankel[2,8,14] & Nevan J. Krogan[1,2,4,14]

**Human immunodeficiency virus (HIV) has a small genome and therefore relies heavily on the host cellular machinery to replicate. Identifying which host proteins and complexes come into physical contact with the viral proteins is crucial for a comprehensive understanding of how HIV rewires the host's cellular machinery during the course of infection. Here we report the use of affinity tagging and purification mass spectrometry[1–3] to determine systematically the physical interactions of all 18 HIV-1 proteins and polyproteins with host proteins in two different human cell lines (HEK293 and Jurkat). Using a quantitative scoring system that we call MiST, we identified with high confidence 497 HIV–human protein–protein interactions involving 435 individual human proteins, with ∼40% of the interactions being identified in both cell types. We found that the host proteins hijacked by HIV, especially those found interacting in both cell types, are highly conserved across primates. We uncovered a number of host complexes targeted by viral proteins, including the finding that HIV protease cleaves eIF3d, a subunit of eukaryotic translation initiation factor 3. This host protein is one of eleven identified in this analysis that act to inhibit HIV replication. This data set facilitates a more comprehensive and detailed understanding of how the host machinery is manipulated during the course of HIV infection.**

A map of the physical interactions between proteins within a particular system is necessary for studying the molecular mechanisms that underlie the system. The analysis of protein–protein interactions (PPIs) has been successfully accomplished in different organisms using a variety of technologies, including mass spectrometry approaches[1,3,4] and those designed to detect pairwise physical interactions, including the two-hybrid yeast system[5,6] and protein-fragment complementation assays[7]. Although two-hybrid methodologies have been used to systematically study host–pathogen interactions[8,9], so far no systematic affinity tagging/purification mass spectrometry (AP–MS) study has been carried out on any host–pathogen system. Here we have targeted HIV-1 for such an analysis, uncovering a wide variety of host proteins, complexes and pathways that are hijacked by the virus during the course of infection.

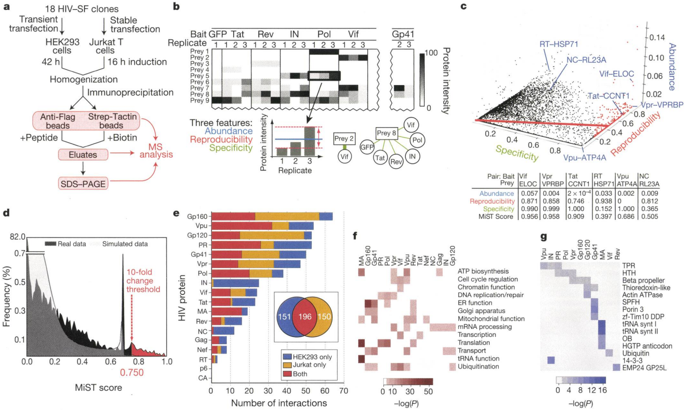
We aimed to identify host proteins associated with HIV-1 proteins systematically and quantitatively using an AP–MS approach[2,3]. To this end, we cloned the genes corresponding to all 18 HIV-1 proteins and polyproteins, including the accessory factors (Vif, Vpu, Vpr and Nef), Tat, Rev, the polyproteins (Gag, Pol and Gp160) and the corresponding processed products (MA, CA, NC and p6; PR, RT and IN; and Gp120 and Gp41, respectively) (Supplementary Fig. 1 and Supplementary Table 1). Each clone was fused to a purification tag (consisting of 2×Strep and 3×Flag) and transiently transfected into HEK293 cells; each also was used to generate stably expressed, tetracycline-inducible, affinity-tagged versions of the proteins in Jurkat cells (Fig. 1a and Supplementary Fig. 2). Following multiple purifications of each factor from both cell lines, the material on the anti-FLAG or Strep-Tactin beads, as well as the eluted material, was analysed by mass spectrometry (Fig. 1a and Supplementary Table 2). Finally, an aliquot of each purified factor was subjected to SDS–polyacrylamide gel electrophoresis, stained (Supplementary Fig. 3) and subjected to analysis by mass spectrometry.

For each HIV factor, we identified co-purifying host proteins that were reproducible regardless of the protocol used (Supplementary Figs 4, 5 and 7 and Supplementary Data 1). Several scoring systems can quantify PPIs from AP–MS proteomic data sets, including PE[10], CompPASS[4] and SAINT[11]. For this data set, we devised a scoring system particularly suited for identifying AP–MS-derived host–pathogen PPIs, which we call MiST (mass spectrometry interaction statistics). The MiST score is a weighted sum of three measures: protein abundance measured by peak intensities from the mass spectrum (abundance); invariability of abundance over replicated experiments (reproducibility); and uniqueness of an observed host–pathogen interaction across all viral purifications (specificity) (Fig. 1b and Supplementary Methods). These three metrics are summed by principal component analysis into a composite score (Fig. 1c and Supplementary Data 2). By comparing our dataset with a benchmark of well-characterized HIV–human PPIs (Supplementary Table 3), analysis of the MiST scoring system revealed superior performance on our data set when compared to CompPASS or SAINT (Supplementary Fig. 6) (and comparable performance using other data sets (Supplementary Fig. 8)) and allowed us to define a MiST cut-off of 0.75, corresponding to ∼4% of all detected interactions. To estimate how many interactions would exceed this threshold by chance, we randomly shuffled our data set 1,000 times. A random MiST score of 0.75 or greater was assigned to an interaction ten times less frequently than we saw among the MiST scores for the real data, and the probability of an interaction assignment with a random MiST score greater than 0.75 was $2.5 \times 10^{-4}$ (Fig. 1d).

At the MiST threshold of 0.75, the number of host proteins we found associated with each HIV protein ranged from 0 (CA and p6) to 63 (Gp160) (Fig. 1e). In total, we observed 497 different HIV–human PPIs (347 and 348 identified from HEK293 cells and Jurkat cells, respectively) (Supplementary Data 3). We detected 196 interactions (∼40%) in both cell types; 150 and 151 were specific to the HEK293 cells and the Jurkat cells, respectively (Fig. 1e). Only some of these

---

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, USA. [2]California Institute for Quantitative Biosciences, QB3, San Francisco, California 94158, USA. [3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, California 94158, USA. [4]J. David Gladstone Institutes, San Francisco, California 94158, USA. [5]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158, USA. [6]Department of Biochemistry, University of Utah, Salt Lake City, Utah 84112, USA. [7]Sanford-Burnham Medical Research Institute, La Jolla, California 92037, USA. [8]Department of Biochemistry and Biophysics, University of California, San Francisco, California 94158, USA. [9]Department of Microbiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [10]The Salk Institute for Biological Studies, La Jolla, California 92037, USA. [11]Department of Chemistry, University of California, Berkeley, California 94720, USA. [12]Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA. [13]Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland. [14]Host Pathogen Circuitry Group, University of California, San Francisco, California 94158, USA. †Present address: Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas, USA, 75390



| Pair: Bait | Vif | Vpr | Tat | RT | Vpu | NC |
|---|---|---|---|---|---|---|
| Prey | ELOC | VPRBP | CCNT1 | HSP71 | ATP4A | RL23A |
| Abundance | 0.057 | 0.004 | $2\times10^{-4}$ | 0.033 | 0.002 | 0.009 |
| Reproducibility | 0.871 | 0.858 | 0.746 | 0.938 | 0 | 0.812 |
| Specificity | 0.990 | 0.999 | 1.000 | 0.152 | 1.000 | 0.365 |
| MiST Score | 0.956 | 0.958 | 0.909 | 0.397 | 0.686 | 0.505 |

**Figure 1 | Affinity purification of HIV-1 proteins, analysis and scoring of mass spectrometry data.** **a**, Flowchart of the proteomic AP–MS used to define the HIV–host interactome. PAGE, polyacrylamide gel electrophoresis. SF, 2×Strep-3×Flag affinity tag. **b**, Data from AP–MS experiments are organized in an interaction table with cells representing amount of prey protein purified (for example spectral counts or peptide intensities). Three features are used to describe bait–prey relationships: abundance (blue), reproducibility (the invariability of bait–prey pair quantities; red) and specificity (green). **c**, All bait–prey pairs are mapped into the three-feature space (abundance, reproducibility and specificity). The MiST score is defined as a projection on the first principal component (red line). All interactions, represented as nodes, above the defined threshold (0.75) are shown in red. This procedure separates the interactions more likely to be biologically relevant (for example Vif–ELOC (ELOC also known as TCEB1), Vpr–VPRBP and Tat–CCNT1) from the interactions that are likely to be less relevant owing to low reproducibility (Vpu–ATP4A) or specificity (RT–HSP71 (HSP71 also known as HSPA1A) and NC–RL23A (RL23A also known as RPL23A)). **d**, The histogram of MiST scores (real data) is compared with a randomized set of scores obtained from randomly shuffling the bait–prey table (simulated data). The MiST score threshold (0.75) was defined using a benchmark (Supplementary Table 3) whereby the predictions are enriched for these interactions by a factor of at least ten relative to random predictions (as well as through ROC (receiver operating characteristic) and recall plots (Supplementary Fig. 6)). **e**, Bar graph of the number of host proteins we found interacting with each HIV factor (MiST score, >0.75). The cell type in which the interaction was found is represented in blue (HEK293 only), yellow (Jurkat only) or red (both). **f**, **g**, Heat maps representing enriched biological functions (**f**) and domains (**g**) from the human proteins identified as interacting with HIV proteins (Supplementary Methods). ER, endoplasmic reticulum; mRNA, messenger RNA; tRNA, transfer RNA. TPR, tetratricopeptide repeat; HTH, helix-turn-helix; SPFH, stomatin–prohibitin–flotillin–HflK/C.

specificities could be explained by differential gene expression in the two cell lines (Supplementary Fig. 9). Using antibodies against 26 of the human proteins, and affinity-tagged versions of an additional 101, we could confirm 97 of the 127 AP–MS derived HIV–human PPIs using co-immunoprecipitation/western blot analysis (76% success rate) (Supplementary Figs 10 and 11), suggesting that we derived a high-quality physical interaction data set.

We next analysed the functional categories of host proteins associated with each HIV protein, and in doing so uncovered many expected connections. These included an enrichment of host factors involved in transcription physically linked to the HIV transcription factor Tat and an enrichment of host machinery implicated in the regulation of ubiquitination associating with Vpu, Vpr and Vif, HIV accessory factors that hijack ubiquitin ligases[12] (Fig. 1f and Supplementary Data 4). When we considered domain types instead of whole proteins (Fig. 1g and Supplementary Table 4), we found that host proteins interacting with IN are enriched in 14-3-3 domains, which generally bind phosphorylated regions of proteins[13], and that proteins containing β-propellers have a higher propensity for binding to Vpr (for additional domain enrichment analysis, see Supplementary Fig. 12). These domain analyses could facilitate future structural modelling of HIV–human PPIs.

Next we compared our data to other HIV-related data sets, including previously published HIV–human PPIs and host factors implicated in HIV function from genome-wide RNA interference (RNAi) screens. For example, the VirusMint database[14] contains 587 HIV–human literature-curated PPIs (Supplementary Data 5), which are mostly derived from small-scale, targeted studies. Although the overlap between the 497 interactions identified in this work and those in VirusMint is statistically significant ($P = 8 \times 10^{-8}$), it corresponds to only 19 PPIs (Fig. 2a and Supplementary Table 5). However, a greater overlap exists, one that remains statistically significant, when interactions below the MiST threshold of 0.75 are considered using a sliding cut-off (for example, at a MiST score of 0.2 there exists an overlap of 67 PPIs ($P = 1 \times 10^{-3}$); Fig. 2c, red lines, and Supplementary Data 6). This overlap indicates that we have indeed identified many interactions that have been previously reported. However, it is likely that the higher scoring interactions identified here have a greater chance of being biologically relevant with respect to HIV function than do many of those in VirusMint.

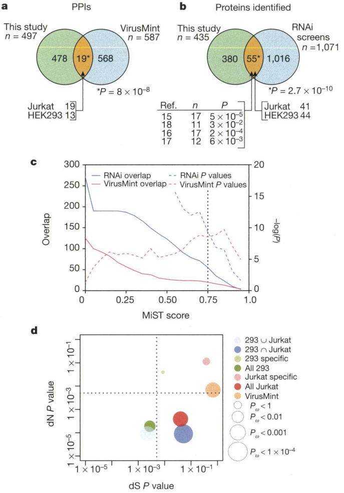


**Figure 2 | Comparison of PPI data with other HIV data sets. a**, Overlap of the 497 HIV–human PPIs with the 587 PPIs reported in VirusMint (Supplementary Table 5). **b**, Overlap of the 435 human proteins with the genes identified in four HIV-dependency RNAi screens[15–18] (Supplementary Table 6). **c**, Number of interactions overlapping with VirusMint (solid red line) and proteins with RNAi screens (solid blue line) as functions of the MiST cut-off. The $P$ values of the overlap are represented as dashed lines using the same colours (Supplementary Data 6 and 8). **d**, Comparative genomics analysis of divergence patterns between human and rhesus macaque reveals strong evolutionary constraints on human proteins binding to HIV proteins. The $x$ and $y$ axes represent $P$ values for the synonymous (dS) and non-synonymous (dN) rates of evolution (Supplementary Methods). Horizontal and vertical dotted lines are drawn at 0.5% to indicate the Bonferroni significance threshold for each axis. For the VirusMint data, the significance of $\omega$ (dN/dS) is primarily driven by higher rates of synonymous evolution. $\cup$, union; $\cap$, intersection; $P_\omega$, bootstrap-based $P$ value for $\omega$.

Recently, four RNAi screens identified host factors that have an adverse effect on HIV-1 replication when knocked down[15–18]. In total, 1,071 human genes were identified in these four studies (Supplementary Data 7), 55 of which overlap with the 435 proteins ($P = 2.7 \times 10^{-10}$; Fig. 2b, Supplementary Fig. 12 and Supplementary Table 6). Again, this overlap increases (as does its statistical significance) if we consider proteins participating in HIV–human PPIs with MiST scores below 0.75 (Fig. 2c, blue lines, and Supplementary Data 8).

To identify the evolutionary forces operating on host proteins interacting with HIV-1, we performed a comparative genomics analysis of divergence patterns between human and rhesus macaque. The proteins identified in both HEK293 and Jurkat cell lines had stronger signatures of evolutionary constraint than those identified exclusively in one cell line or in VirusMint (Fig. 2d). Points in the lower-right quadrant of Fig. 2d show signatures of strong purifying selection, whereas the upper-right quadrant shows signatures more consistent with neutral evolution. This observation suggests that the PPIs identified in our study, especially the ones identified in both cell types, are more physiologically relevant to mammalian evolution than those reported in VirusMint.

We next plotted the 497 HIV–human interactions identified in this study in a network representation (Fig. 3) containing nodes corresponding to 16 HIV (yellow) and 435 human factors that were derived from the HEK293 cells (blue), Jurkat cells (red) or both. We also introduced 289 interactions between human proteins (black edges) derived from several databases[19,20] (Supplementary Data 9). These human–human interactions helped to identify many host complexes, including several that have been previously characterized (see Supplementary Information for a detailed discussion of the HIV–human interaction data sets). Ultimately, all data will be accessible for searching and comparison to other HIV-related data sets using the web-based software GPS-PROT[21] (http://www.gpsprot.org/).

Notably, we found that Pol and PR, which we needed to make catalytically inactive (Supplementary Fig. 1), bound the translational initiation complex eIF3, a 13-subunit complex (eIF3a to eIF3m). We detected 12 of the subunits bound to Pol and/or PR, except eIF3j, which is only loosely associated with the complex[22] (Fig. 4a). Even though PR is the smallest of the *pol*-encoded proteins, we find it associated with the greatest number of host factors (Fig. 4a). To determine whether components of the translation complex are substrates for PR, FLAG-tagged versions of ten eIF3 subunits were individually co-transfected, each with a small amount of active HIV-1 PR, into HEK293 cells. The cell lysates were analysed by western blotting and only eIF3d was found to be cleaved (Fig. 4b). Purification of tagged versions of the amino and carboxy termini of cleaved eIF3d revealed that only the N terminus of 114 amino-acid residues associates with the eIF3 complex (Supplementary Table 7). The cleavage occurred with an efficiency similar to that of the processing of the natural PR substrate Gag (Fig. 4c), whereas two cellular proteins previously described to be cleaved by HIV PR, PAPBC1[23] and BCL2[24], were cleaved only at higher PR concentrations or not at all, respectively. To confirm this result *in vitro*, we incubated purified human eIF3 with active PR, resulting in the removal of a 70-kDa band and the appearance of a ~60-kDa protein product (Fig. 4d). Analysis of the cleaved product by N-terminal sequencing revealed a cleavage of eIF3d between Met 114 and Leu 115, which corresponds to the consensus sequence for HIV-1 protease[25] and falls within the RNA-binding domain (RRM) of eIF3d (ref. 26; Fig. 4d).

Next we used four to six short interfering RNAs against different eIF3 subunits in HIV infectivity assays (Fig. 4e, f, Supplementary Fig. 14 and Supplementary Table 8). Using a fusion of HIV with vesicular stomatitis virus glycoprotein (VSV-G), which only allows for a single round of replication, knockdown of eIF3d, but not other eIF3 subunits, resulted in an increase in infectivity (Fig. 4e), suggesting that this factor acts in early stages of infection. In assays requiring multiple rounds of HIV infection, knockdown of eIF3d, eIF3e and eIF3f enhanced HIV NL4.3 infectivity by a factor of three to five, whereas inhibition of eIF3c, eIF3g and eIF3i had no promoting effect (Fig. 4f). Consistent with these results, a previous overexpression screen for factors that restrict HIV-1 replication identified eIF3f as the most potent inhibitory clone[27]. Furthermore, using assays monitoring both early and late products we found that knockdown of eIF3d results in an increase in accumulation of reverse transcription product (Fig. 4g and Supplementary Fig. 15). This suggests that eIF3 does in fact have a role in the early stages of infection, perhaps by binding to the viral RNA through the RNA-binding domain in eIF3d, and thus inhibiting RT, an effect that is overcome by PR cleavage of eIF3d (Supplementary Fig. 16). These results suggest that our data set will be enriched not only for host proteins the virus requires for efficient replication (Fig. 2b, c), but also those that have an inhibitory role during infection. Indeed, we have found that an additional ten factors from our list of interactors, when knocked down by RNAi, produce an increase in HIV infection (Supplementary Figs 17–19, Supplementary Tables 12

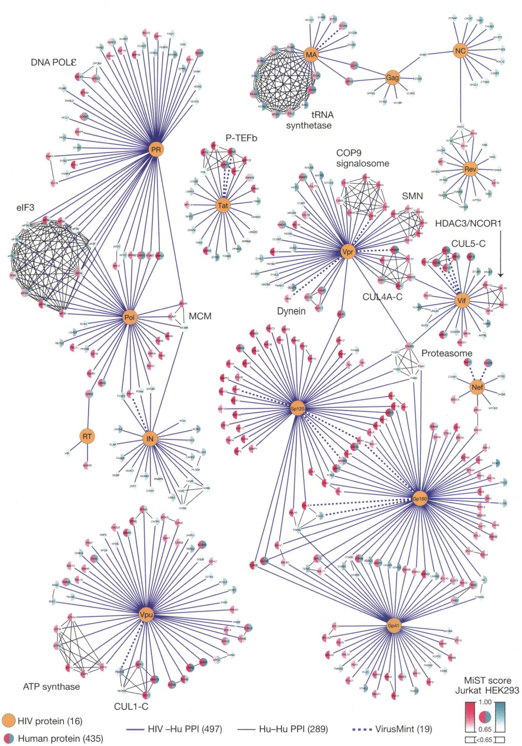
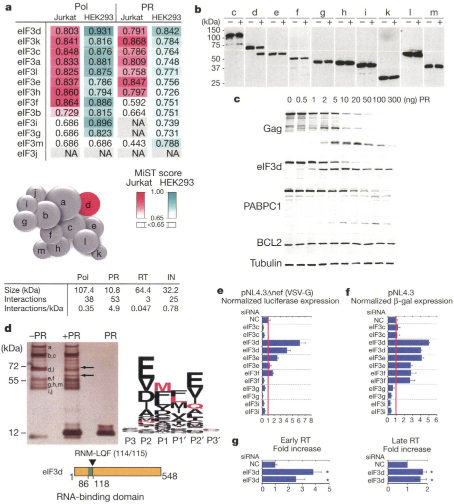

DNA POLℇ
PR
eIF3
Pol
MCM
RT
IN
P-TEFb
Tat
tRNA
synthetase
MA
Gag
NC
Rev
COP9
signalosome
SMN
HDAC3/NCOR1
CUL5-C
Vpr
Vif
CUL4A-C
Dynein
Proteasome
Nef
Gp120
Gp160
Gp41
Vpu
ATP synthase
CUL1-C
HIV protein (16)
Human protein (435)
HIV -Hu PPI (497)
Hu-Hu PPI (289)
VirusMint (19)
MiST score
Jurkat HEK293
1.00
0.65
<0.65

**Figure 3 | Network representation of the HIV–human PPIs.** In total, 497 HIV–human interactions (blue) are represented between 16 HIV proteins and 435 human factors. Each node representing a human protein is split into two colours and the intensity of each colour corresponds to the MiST score from interactions derived from HEK293 (blue) or Jurkat (red) cells. Black edges correspond to interactions between host factors (289) that were obtained from publicly available databases; dashed edges correspond to interactions also found in VirusMint[14].

and 13 and Supplementary Methods). Knockdown of two of these, DESP and HEAT1, also resulted in an increase in HIV integration (Supplementary Fig. 20 and Supplementary Table 14), consistent with their physical association with IN.

As well as performing the systematic AP–MS study reported here, we explored in further detail the biological significance of two newly identified HIV–human interactions: HIV protease targeting a component of eIF3 that is inhibitory to HIV replication; and CBF-β, a new component of the Vif–CUL5 ubiquitin ligase complex required for APOBEC3G stability and HIV infectivity[28]. Further work will be required to determine whether, how and at what stage of infection the remaining host factors impinge on HIV function. Ultimately, our analysis of the host factors co-opted by different viruses using the same proteomic pipeline will allow for the identification of protein complexes routinely targeted by different pathogens, which may represent better therapeutic targets for future studies.



| | Pol Jurkat | Pol HEK293 | PR Jurkat | PR HEK293 |
|---|---|---|---|---|
| eIF3d | 0.803 | 0.931 | 0.791 | 0.842 |
| eIF3k | 0.841 | 0.816 | 0.868 | 0.784 |
| eIF3c | 0.848 | 0.876 | 0.786 | 0.764 |
| eIF3a | 0.833 | 0.881 | 0.809 | 0.748 |
| eIF3l | 0.825 | 0.875 | 0.758 | 0.771 |
| eIF3e | 0.837 | 0.786 | 0.847 | 0.756 |
| eIF3h | 0.860 | 0.794 | 0.797 | 0.726 |
| eIF3f | 0.864 | 0.886 | 0.592 | 0.751 |
| eIF3b | 0.729 | 0.815 | 0.664 | 0.751 |
| eIF3i | 0.686 | 0.896 | NA | 0.739 |
| eIF3g | 0.686 | 0.823 | NA | 0.731 |
| eIF3m | 0.686 | 0.686 | 0.443 | 0.788 |
| eIF3j | NA | NA | NA | NA |

| | Pol | PR | RT | IN |
|---|---|---|---|---|
| Size (kDa) | 107.4 | 10.8 | 64.4 | 32.2 |
| Interactions | 38 | 53 | 3 | 25 |
| Interactions/kDa | 0.35 | 4.9 | 0.047 | 0.78 |

a
MiST score
Jurkat HEK293
1.00
0.65
<0.65
j i a d b g c e f l h m k
b
(kDa)
150
100
75
50
37
25
c d e f g h i k l m
− + − + − + − + − + − + − + − + − + − +
c
0 0.5 1 2 5 10 20 50 100 300 (ng) PR
Gag
eIF3d
PABPC1
BCL2
Tubulin
d
−PR +PR PR
(kDa)
72
55
12
a
b,c
d,l
e,f
g,h,m
i,j
P3 P2 P1 P1′ P2′ P3′
RNM-LQF (114/115)
eIF3d
1
86 118
548
RNA-binding domain
e
pNL4.3Δnef (VSV-G)
Normalized luciferase expression
f
pNL4.3
Normalized β-gal expression
siRNA
NC
eIF3c
eIF3c
eIF3d
eIF3d
eIF3e
eIF3e
eIF3f
eIF3f
eIF3g
eIF3g
eIF3i
eIF3i
0 1 2 3 4 5 6 7 8
0 1 2 3 4 5 6
g
Early RT
Fold increase
Late RT
Fold increase
siRNA
NC
eIF3d
eIF3d
*
0 1 2 3 4 5
0 1 2

**Figure 4 | eIF3d is cleaved by HIV-1 PR and inhibits infection. a,** MiST scores for eIF3 subunits associated with PR (right) and Pol (left) in HEK293 and Jurkat cells. Sizes of the proteins and numbers of significant interactions (MiST score, >0.75) detected for Pol and its subunits are shown below, as is a modular representation of the eIF3 complex[29]. The cleaved subunit, eIF3d, is in red. **b,** Western blot of HEK293 cell lysate expressing FLAG-tagged eIF3 subunits in the absence (−) or presence (+) of active PR probed with an anti-FLAG antibody. **c,** HEK293 cells were co-transfected with Gag, or FLAG-tagged eIF3d, PABPC1, BCL2 and increasing amounts of PR. Cell lysates were probed against Gag (upper panel), FLAG-tagged eIF3d (middle panel) or tubulin as control (lower panel). **d,** Silver stain of purified eIF3 complex incubated with recombinant HIV-1 PR. The residues corresponding to the eIF3d cleavage site (red) is located within the RNA-binding domain[26]. **e, f,** HeLa-derived P4/R5 MAGI cells were transfected with two different short interfering RNAs (siRNAs) targeting individual subunits of the eIF3 complex (Supplementary Tables 7 and 9) and subsequently infected with either a pNL4-3-derived, VSV-G-pseudotyped, single-cycle virus (HIV–VSV-G) (**e**) or wild-type pNL4-3 (**f**). NC, negative control. **g,** Early (left) and late (right) HIV-1 DNA levels measured by quantitative PCR amplification in cells transfected with two independent eIF3d siRNAs or with control siRNAs. Samples were normalized by input DNA amount or by cellular gene (*HMBS*) copy number. The RT and replication assays were done three to five times and the standard deviations are shown (Supplementary Tables 7, 10 and 11). $^*P < 0.05$ (Kruskal–Wallis test with Dunn's correction for multiple comparisons).

## METHODS SUMMARY

More details on experimental assays, plasmid constructs, sequences, cell lines, antibodies and computational analysis are provided in Supplementary Methods. Briefly, affinity tagging and purification was carried out as previously described[2] and the protein samples were analysed on a Thermo Scientific LTQ Orbitrap XL mass spectrometer. For the evolutionary analysis, genome-wide alignments to rhesus macaque were downloaded from the University of California, Santa Cruz genome browser (http://genome.ucsc.edu/) and evolutionary rates for each group of genes considered were measured using the synonymous and non-synonymous rates of evolution. For the *in vitro* protease assay, maltose binding protein (MBP)-tagged PR was expressed in BL21 (Gold) DE3 cells in the presence of 100 μM Saquinavir and purified on an MBP trap column. Purified eIF3 was obtained from J. Cate (University of California, Berkeley). For the infection assays, HeLa P4.R5 cells were transfected with short interfering RNAs and after 48 h infected with pNL4-3 or a pNL4-3-derived VSV-G-pseudotyped reporter virus. Infection levels were determined by luminescence read-out.

**Received 26 March; accepted 18 November 2011.**

**Published online 21 December 2011.**

1. Gavin, A. C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440,** 631–636 (2006).
2. Jäger, S. *et al.* Purification and characterization of HIV-human protein complexes. *Methods* **53,** 13–19 (2011).
3. Krogan, N. J. *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae. Nature* **440,** 637–643 (2006).
4. Sowa, M. E., Bennett, E. J., Gygi, S. P. & Harper, J. W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138,** 389–403 (2009).
5. Yu, H. *et al.* High-quality binary protein interaction map of the yeast interactome network. *Science* **322,** 104–110 (2008).
6. Stelzl, U. *et al.* A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122,** 957–968 (2005).
7. Tarassov, K. *et al.* An *in vivo* map of the yeast protein interactome. *Science* **320,** 1465–1470 (2008).
8. Calderwood, M. A. *et al.* Epstein-Barr virus and virus human protein interaction maps. *Proc. Natl Acad. Sci. USA* **104,** 7606–7611 (2007).

9. Shapira, S. D. *et al.* A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139,** 1255–1267 (2009).
10. Collins, S. R. *et al.* Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae. Mol. Cell. Proteomics* **6,** 439–450 (2007).
11. Choi, H. *et al.* SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nature Methods* **8,** 70–73 (2011).
12. Malim, M. H. & Emerman, M. HIV-1 accessory proteins–ensuring viral survival in a hostile environment. *Cell Host Microbe* **3,** 388–398 (2008).
13. Yaffe, M. B. *et al.* The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91,** 961–971 (1997).
14. Chatr-aryamontri, A. *et al.* VirusMINT: a viral protein interaction database. *Nucleic Acids Res.* **37,** D669–D673 (2009).
15. Brass, A. L. *et al.* Identification of host proteins required for HIV infection through a functional genomic screen. *Science* **319,** 921–926 (2008).
16. König, R. *et al.* Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **135,** 49–60 (2008).
17. Yeung, M. L., Houzet, L., Yedavalli, V. S. & Jeang, K. T. A genome-wide short hairpin RNA screening of Jurkat T-cells for human proteins contributing to productive HIV-1 replication. *J. Biol. Chem.* **284,** 19463–19473 (2009).
18. Zhou, H. *et al.* Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* **4,** 495–504 (2008).
19. Ruepp, A. *et al.* CORUM: the comprehensive resource of mammalian protein complexes–2009. *Nucleic Acids Res.* **38,** D497–D501 (2010).
20. Stark, C. *et al.* BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34,** D535–D539 (2006).
21. Fahey, M. E. *et al.* GPS-Prot: a web-based visualization platform for integrating host-pathogen interaction data. *BMC Bioinformatics* **12,** 298 (2011).
22. Hinnebusch, A. G. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem. Sci.* **31,** 553–562 (2006).
23. Alvarez, E., Castello, A., Menendez-Arias, L. & Carrasco, L. HIV protease cleaves poly(A)-binding protein. *Biochem. J.* **396,** 219–226 (2006).
24. Strack, P. R. *et al.* Apoptosis mediated by HIV protease is preceded by cleavage of Bcl-2. *Proc. Natl Acad. Sci. USA* **93,** 9571–9576 (1996).
25. Schilling, O. & Overall, C. M. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nature Biotechnol.* **26,** 685–694 (2008).
26. Asano, K. *et al.* Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. *J. Biol. Chem.* **272,** 27042–27052 (1997).
27. Valente, S. T., Gilmartin, G. M., Mott, C., Falkard, B. & Goff, S. P. Inhibition of HIV-1 replication by eIF3f. *Proc. Natl Acad. Sci. USA* **106,** 4071–4078 (2009).
28. Jäger, S. *et al.* Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. *Nature* http://dx.doi.org/10.1038/nature10693 (this issue).
29. Zhou, M. *et al.* Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proc. Natl Acad. Sci. USA* **105,** 18139–18144 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank A. Choi, Z. Rizvi and E. Kwon for cloning of human genes and J. Cate for purified eIF3. We also thank J. Gross, R. Andino, R. Harris, M. Daugherty and members of the Krogan lab for discussion. This research was funded by grants from QB3@UCSF and the National Institutes of Health (P50 GM082250 to N.J.K., A.D.F., C.S.C. and T.A.; P01 AI090935 to N.J.K., S.K.C., J.A.Y. and F.D.B.; P50 GM081879 to N.J.K. and A.B.; P50 GM082545 to W.I.S.; P41RR001614 to A.B.; U54 RR022220 to A.S.; P01 GM073732-05 to A.T.; CHRP-ID08-TBI-063 to S.K.C.; P41 RR001081 to J.H.M.) and from the Nomis Foundation (to J.A.Y.). N.J.K. is a Searle Scholar and a Keck Young Investigator.

**Author Contributions** S.J. generated the protein–protein interaction map; P.C. developed the MiST scoring system; N.G., M. Shales, E.A., M.F., J.H.M., J.R.J. and R.D.H. provided computational support; K.E.M., K.L., J.R.J., H.H., G.M.J., I.D., J.F. and D.A.M. provided experimental support; S.J., S.C.C., A.J.O. and A.T. characterized the PR–eIF3d interaction; S.J., G.M.J., C.M. and G.M. confirmed the interactions by immunoprecipitation/western blot; L.P., S.L.R., J.M. and M. Stephens used RNAi for functional verification; T.A., G.C., F.D.B., J.A.Y., S.K.C., W.I.S., T.K., R.D.H., C.S.C., A.B., A.S., A.D.F. and N.J.K. supervised the research; and S.J., P.C., A.S. and N.J.K. wrote the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.J.K. (krogan@cmp.ucsf.edu).

LETTER

doi:10.1038/nature10693

# Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection

Stefanie Jäger[1,2]*, Dong Young Kim[3]*, Judd F. Hultquist[4]*, Keisuke Shindo[4], Rebecca S. LaRue[4], Eunju Kwon[3], Ming Li[4], Brett D. Anderson[4], Linda Yen[3], David Stanley[3], Cathal Mahon[1,2,3], Joshua Kane[1,2], Kathy Franks-Skiba[1,2], Peter Cimermancic[2,5], Alma Burlingame[2,3], Andrej Sali[2,3,5,6], Charles S. Craik[2,3], Reuben S. Harris[4], John D. Gross[2,3,6] & Nevan J. Krogan[1,2,6,7]

**Restriction factors, such as the retroviral complementary DNA deaminase APOBEC3G, are cellular proteins that dominantly block virus replication[1–3]. The AIDS virus, human immunodeficiency virus type 1 (HIV-1), produces the accessory factor Vif, which counteracts the host's antiviral defence by hijacking a ubiquitin ligase complex, containing CUL5, ELOC, ELOB and a RING-box protein, and targeting APOBEC3G for degradation[4–10]. Here we reveal, using an affinity tag/purification mass spectrometry approach, that Vif additionally recruits the transcription cofactor CBF-β to this ubiquitin ligase complex. CBF-β, which normally functions in concert with RUNX DNA binding proteins, allows the reconstitution of a recombinant six-protein assembly that elicits specific polyubiquitination activity with APOBEC3G, but not the related deaminase APOBEC3A. Using RNA knockdown and genetic complementation studies, we also demonstrate that CBF-β is required for Vif-mediated degradation of APOBEC3G and therefore for preserving HIV-1 infectivity. Finally, simian immunodeficiency virus (SIV) Vif also binds to and requires CBF-β to degrade rhesus macaque APOBEC3G, indicating functional conservation. Methods of disrupting the CBF-β–Vif interaction might enable HIV-1 restriction and provide a supplement to current antiviral therapies that primarily target viral proteins.**

Mammals have evolved cellular proteins termed restriction factors that function to prevent the spread of mobile genetic elements including retroviruses[1–3]. As a counter-defence, most retroviruses, including the human pathogen HIV-1, have developed mechanisms to prevent restriction, often through subversion of the host's ubiquitin–proteasome system. In eukaryotic cells, 8.6-kDa ubiquitin moieties are added to a target protein by sequential action of one of two ubiquitin-activating enzymes (E1), which transfer ubiquitin to a pool of dozens of ubiquitin-conjugating enzymes (E2) that, in turn, collaborate with hundreds of ubiquitin ligases (E3) to catalyse transfer to specific substrates[11]. If more than four ubiquitins are joined together through K48 linkages, the target protein is usually degraded by the 26S proteasome[12]. At least three HIV-1 proteins, Vif, Vpu and Vpr, hijack cullin-RING E3 ligases consisting of CUL5, CUL1 and CUL4A to promote ubiquitination and degradation of APOBEC3 family members (for example, APOBEC3G, A3G), BST2/tetherin and an unknown, putative restriction factor, respectively[2]. Understanding the composition of cullin-RING E3 ligase complexes and the underlying cellular signalling components may provide therapeutic routes for treating a variety of human diseases, including infection by HIV-1.

HIV-1 Vif is recruited to CUL5 by virtue of its SOCS box, which contains an elongin C binding helix (the BC-box), a conserved HCCH Zn binding motif and a short Cullin Box[4–6]. Although a structure of the BC-box peptide in complex with the heterodimer of Elongin B and C (ELOBC) has been reported[13], the architecture of the full-length Vif in complex with host factors has remained elusive, in part because Vif complexes have poor solubility and activity. We therefore reasoned that Vif may bind an additional host factor and that such a factor may render it more tractable *in vitro*.

We took an unbiased proteomic approach to identify host factors that bind all 18 HIV processed and polyproteins using an affinity tag/purification mass spectrometry (AP–MS) approach[14,15]. To this end, 2×Strep and 3×Flag was fused to the carboxy (C) terminus of these factors, including Vif. The tagged Vif construct was both transiently transfected into HEK293 cells and used to make a stable, tetracycline-inducible Vif–Strep–Flag Jurkat T cell line (Fig. 1a). Epitope-tagged Vif was purified from both cell types using antibodies specific to either Strep or Flag and aliquots of the co-purifying proteins were subjected to SDS-polyacrylamide gel electrophoresis (SDS–PAGE) (Fig. 1b). Materials from each step were analysed by mass spectrometry[14].

Using a new scoring system for data derived from AP–MS studies, termed Mass Spectrometry Interaction Statistics (MiST)[15], we identified 24 Vif–human protein–protein interactions with seven of them found in both cell types (Fig. 1c). Seventeen of these were verified independently by co-immunoprecipitation (Supplementary Fig. 1). Among these were the components of the E3 ubiquitin ligase complex, CUL5, ELOB and ELOC, known to interact with Vif and trigger A3G degradation[4–6,8–10]. Although the RING-box protein RBX1 was originally reported as part of this complex[4], only RBX2 was above the MiST score threshold used[15] consistent with recent work showing that it binds CUL5 (refs 16, 17). We did not find endogenous A3G, probably because of its poor expression in HEK293 and Jurkat cell lines exacerbated by further depletion through Vif-mediated degradation. We did find Vif associating with two proteins that function in autophagy, AMRA1 and SQSTM, as well as with the transcriptional co-repressor complex NCOR1/HDAC3/GPS2/TBL1R (the last only in T cells) (Fig. 1c). Also, in both cell types, Vif was found to interact with the transcription cofactor CBF-β, which is known to heterodimerize with the RUNX family of transcription factors[18].

To determine if any of the newly defined Vif interactors belong to the Vif–CUL5 complex, we performed double affinity purifications using cells co-transfected with Vif–2×Strep and either A3G– or CUL5–3×Flag (Fig. 1d, e). After purification first with Strep-Tactin and second with anti-Flag beads, mass spectrometry analysis of the final elution revealed the presence of CUL5, ELOB, ELOC, RBX2 and invariably CBF-β, strongly suggesting that this last protein may be a new component of the Vif E3 ubiquitin ligase complex (other factors from single Vif purifications depicted in Fig. 1c were not present). To confirm this interaction and the composition of the complex, we performed an additional double affinity purification experiment using

---

[1]Department of Cellular and Molecular Pharmacology, University of California-San Francisco, San Francisco, California 94158, USA. [2]California Institute for Quantitative Biosciences, QB3, San Francisco, California, California 94158, USA. [3]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158, USA. [4]Department of Biochemistry, Molecular Biology and Biophysics, Institute for Molecular Virology, Center for Genome Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA. [5]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, California 94158, USA. [6]HPC (Host Pathogen Circuitry) Group, University of California-San Francisco, San Francisco, California 94158, USA. [7]J. David Gladstone Institutes, San Francisco, California 94158, USA.
*These authors contributed equally to this work.

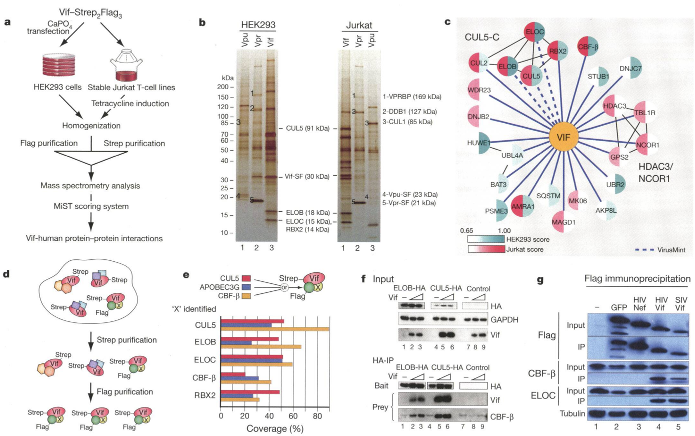


**Figure 1 | AP–MS experiments identify CBF-β as a Vif-dependent component of the Vif–CUL5 ubiquitin ligase complex. a**, Flow-chart of the proteomic analysis performed during the study. **b**, Affinity-tagged versions of Vif, Vpu and Vpr were purified using 3×Flag from HEK293 and Jurkat cells, subjected to SDS–PAGE and stained with silver. Visible bands corresponding to interactions that are known for each accessory factor are labelled. Note Vif and CBF-β run at a similar place on the gel. Tagged versions of Vpr and Vpu were used as specificity controls. **c**, A network representation of Vif–host protein–protein interactions from both HEK293 (blue) and Jurkat T cells (red) after subjecting the data derived from the AP–MS analysis to the MiST scoring system[15]. The intensity of the node colours corresponds to the quantitative MiST score. Blue edges represent interactions derived during this work; black edges are previously described interactions between host factors; dashed edges correspond to previously described Vif–host interactions present in the database VirusMint. **d**, The double purification approach, which allows for the identification of stable, stoichiometric protein complexes. **e**, Double purifications were performed in triplicate using 3×Flag-tagged CUL5, A3G or CBF-β with 2×Strep-tagged Vif in HEK293 cells. Proteins that were identified in all three double purifications, after trypsin digestion and analysis by mass spectrometry, are represented. The coverage corresponds to the percentage of protein identified by tryptic peptides. **f**, Immunoblots showing that Vif recruits CBF-β to the CUL5/ELOBC/RBX2 ubiquitin ligase complex. HA-tagged ELOB or CUL5 were immunoprecipitated in the absence or presence of increasing amounts of Vif, and endogenous CBF-β was monitored by immunoblot. **g**, HIV and SIV Vif co-immunoprecipitate CBF-β and ELOC. GFP and HIV Nef were analysed in parallel as specificity controls.

Vif–2×Strep and CBF-β–3×Flag. This strategy also yielded CUL5, ELOB, ELOC and RBX2, in addition to the epitope-tagged bait proteins (Fig. 1d, e).

To determine if the association of CBF-β with the CUL5 ligase complex was dependent on Vif, we immunoprecipitated CUL5-haemagglutinin (HA) or ELOB-HA in the presence or absence of Vif in HEK293 cells and blotted for endogenous CBF-β. Only in the presence of Vif did CBF-β co-immunoprecipitate with tagged CUL5 or ELOB, indicating that recruitment of CBF-β to the CUL5 ligase is dependent on Vif (Fig. 1f). SIV Vif also associated with CBF-β by immunoprecipitation, suggesting the interaction is conserved (Fig. 1g).

We next asked if the Vif–CUL5 ligase could be reconstituted with CBF-β using recombinant proteins purified from *Escherichia coli*. Initial purification attempts without CBF-β yielded aggregated and inactive complexes, assayed by size-exclusion chromatography and autoubiquitination activity, suggesting that CBF-β may be required for complex formation (data not shown). Therefore, full-length Vif, ELOB, ELOC and CBF-β were co-expressed, purified to homogeneity and found to form a stable, monodisperse complex with recombinant CUL5/RBX2, as shown by size-exclusion chromatography and SDS–PAGE analysis (Fig. 2a, b). Pull-down experiments performed with purified, His-tagged APOBEC3 enzymes immobilized on cobalt-chelating resin showed that the four protein complex containing Vif, CBF-β and ELOBC binds A3G, but not the related Vif-resistant deaminase, A3A (Fig. 2c). These observations suggested that Vif, CBF-β and ELOBC form a substrate adaptor for CUL5/RBX2 that enables specific interaction with susceptible A3 proteins.

To test the activity of the reconstituted six protein complex, CUL5/RBX2/ELOB/ELOC/Vif/CBF-β (CRL5–Vif–CBF-β), we assayed substrate and Vif ubiquitination activities using two distinct and well characterized ubiquitin conjugating enzymes, UBE2R1 (hCDC34a) and UBCH5b, which are capable of forming specific K48 and heterogenous ubiquitin chain linkages, respectively[19,20]. With UBE2R1, CRL5–Vif–CBF-β catalysed formation of high-molecular mass K48 chains on A3G, but not A3A (Fig. 2d, e), mirroring the chain linkage and substrate specificity observed in cells[4,6,21–23]. As with most ubiquitin ligase assemblies, the CRL5–Vif–CBF-β complex also possessed autoubiquitination activity that was only marginally affected by substrate A3s (Supplementary Fig. 2). These experiments were done with NEDD8-modified CUL5, because NEDD8ylation is required for CUL5 to degrade A3G *in vivo*[4] (Supplementary Fig. 3). Similarly, with UBCH5b, CRL5–Vif–CBF-β was able to promote the specific polyubiquitination of A3G and elicit Vif autoubiquitination activity (Supplementary Fig. 4). We conclude that the reconstituted Vif E3 ligase is specific for A3G,

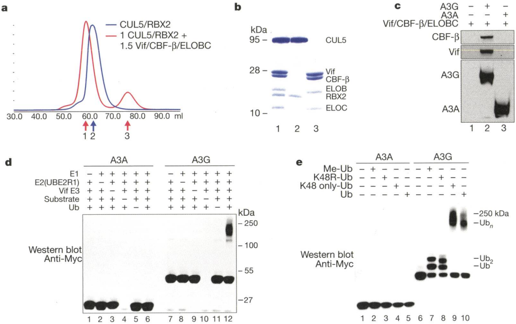


**Figure 2 | CBF-β is a stoichiometric component of the Vif E3 ubiquitin ligase.** **a**, Size exclusion chromatography of recombinant purified CUL5/RBX2 (blue) overlaid with CUL5/RBX2 mixed with 1.5 equivalents of purified Vif substrate adaptor containing Vif, ELOBC and CBF-β (red). **b**, Coomassie-stained SDS–PAGE of fractions labelled 1–3 in **a** indicating the Vif substrate adaptor and a six-protein assembly (CRL5–Vif–CBF-β) co-purify as stable monodisperse species. **c**, A3G, but not A3A, directly binds the tetrameric Vif substrate adaptor in pull-down experiments *in vitro*. **d**, CRL5–Vif–CBF-β is an E3 ligase that promotes polyubiquitination of A3G, but not A3A (detected using an anti-c-Myc antibody to the C-terminal tag on the deaminases). Ub, ubiquitin. **e**, CRL5–Vif–CBF-β and UBE2R1 catalyse formation of K48-linked chains on A3G. Immunoblots showing substrate in ubiquitination reactions containing UBE2R1 as E2, no ubiquitin, Me-ubiquitin, K48R-ubiquitin, K48-only ubiquitin or wild-type ubiquitin. Reactions with Me-ubiquitin indicate at least two distinct sites are modified on A3G; K48R recapitulates the pattern observed with Me-ubiquitin, whereas both wild type and K48R-only ubiquitin result in extensive polyubiquitin chains.

supports K48 chain formation and can function with at least two ubiquitin conjugating enzymes *in vitro*. It is conceivable that these two ubiquitin-conjugating enzymes work together in cells to promote multi-monoubiquitination of A3G followed by specific chain elongation, as described for other RING E3s[24,25], but additional work will be necessary to rule out other E2s *in vivo*.

To determine if CBF-β is required for Vif folding and/or stability in living cells, we transfected a constant amount of Vif into HEK293T cells expressing either a scrambled short hairpin (sh)RNA or a CBF-β-specific shRNA. The levels of steady-state Vif were threefold lower in CBF-β-depleted cells than in the scrambled control cells (Fig. 3a). Proteasome inhibitor MG132 reversed this effect, suggesting that Vif

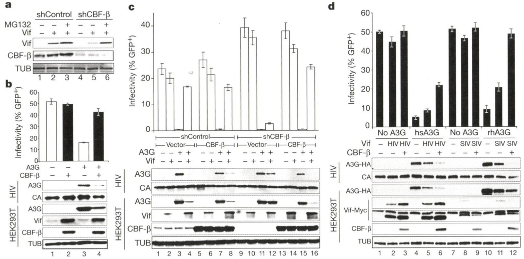


**Figure 3 | CBF-β and Vif collaborate to degrade APOBEC3G and enable HIV-1 infectivity.** **a**, CBF-β-depleted HEK293T cells have lower steady-state Vif levels, which recover upon treatment with 2.5 μM MG132. **b**, Infectivity of replication-competent, Vif-proficient HIV-1 in the presence and absence of CBF-β and A3G ($n = 3$; mean, s.d.). Immunoblots are shown for the indicated proteins in virus-producing cells and viral particles. **c**, Infectivity of HIV-GFP produced using HEK293T-shCBF-β or HEK293T-shControl clones transfected with the single-cycle virus cocktail, A3G, Vif and CBF-β as indicated ($n = 3$; mean, s.d.). The corresponding immunoblots are shown below. **d**, Infectivity of a Vif-deficient HIV-1 molecular clone produced in the presence or absence of human or rhesus A3G-HA, HIV or SIV Vif–Myc, and CBF-β as indicated ($n = 3$; mean, s.d.). Immunoblots are shown for the indicated proteins in virus-producing cells and viral particles with two exposures of the anti-Myc (Vif) blot shown to clarify the SIV Vif signal (the longer exposure also shows endogenous c-Myc).

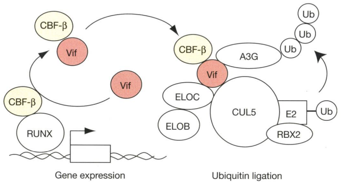


**Figure 4 | Model for Vif–CBF-β E3 ligase formation and APOBEC3G polyubiquitination and degradation.** Vif is depicted hijacking cellular CBF-β to the E3 ubiquitin ligase complex required for A3G polyubiquitination and degradation. Vif may recruit newly translated CBF-β (not shown) and/or hijack existing CBF-β from RUNX transcription complexes.

degradation is accelerated without CBF-β. Analogous data were obtained when Vif was expressed from a proviral plasmid in CBF-β-depleted cells and complemented with a CBF-β expression plasmid (Supplementary Fig. 5).

Based on these observations, we predicted that CBF-β knockdown should result in less functional Vif and less infectious HIV-1 particles when produced in the presence of A3G. To test this prediction, shRNA was used to deplete CBF-β stably in HEK293T cells, and a knockdown clone was used to produce replication competent Vif-proficient HIV-1 in the presence or absence of A3G and CBF-β expressed from plasmids (Fig. 3b). In CBF-β depleted cells, steady-state Vif levels were very low despite equivalent levels of virus production as indicated by capsid. Moreover, Vif levels increased when CBF-β was replenished by complementation, and this correlated with decreases in cellular and viral A3G levels and corresponding increases in viral infectivity. In the absence of A3G, no difference in infectivity was observed regardless of cellular CBF-β or Vif levels. Titration experiments showed that CBF-β complementation is dose-responsive (Supplementary Fig. 6). Analogous results were obtained with a multi-vector HIV/green fluorescent protein (GFP) system (Fig. 3c and Supplementary Fig. 7). The Vif/CBF-β interaction was confirmed in virus-producing cells by co-immunoprecipitation experiments (Supplementary Fig. 8). Furthermore, SIVmac239 Vif requires CBF-β to degrade rhesus macaque A3G and promote viral infectivity (Fig. 3d). Interestingly, in contrast to HIV Vif, lower steady-state levels of SIV Vif were observed in the presence of CBF-β, which may be functionally significant or may be a consequence of the heterologous assay system (that is, expressing SIV/rhesus proteins in human cells). Nevertheless, these results demonstrate the essential and conserved nature of CBF-β for Vif function in promoting A3G degradation and efficient virus replication.

Our proteomic, biochemical and genetic studies combine to suggest a model in which HIV-1 Vif hijacks the cellular transcription factor CBF-β to facilitate Vif folding and/or stability as well as nucleation of the rest of the E3 ubiquitin ligase complex (Fig. 4). CBF-β is required for A3G substrate binding and, ultimately, for polyubiquitination and degradation, thereby enabling the production of infectious viral particles. Because genetic studies have shown that Vif is also capable of degrading APOBEC3F and several other human APOBEC3 proteins[2,3,23] most of which are expressed in primary CD4+ T lymphocytes[26,27], it is quite likely that CBF-β is required for counteracting multiple endogenous APOBEC3s and thus for rendering T lymphocytes permissive for HIV-1 replication. We anticipate that the development of antiviral therapies that antagonize the CBF-β–Vif interaction will be more powerful than those that specifically target the A3G–Vif interaction, because they have the potential to unleash the simultaneous restriction potential of multiple APOBEC3s analogous to current combinatorial therapies.

## METHODS SUMMARY

Affinity tagging, purification[14] and ubiquitination assays[28] were performed as described. Recombinant A3A and A3G were purified as Myc-His tagged proteins from HEK293 cells. HIV-1 infectivity studies used an HIV-1$_{\mathrm{IIIB}}$ proviral DNA construct (with or without Vif) or an HIV-GFP reporter plasmid set. Control (RHS4346) and CBF-β (RHS4430-99161432) shRNA constructs were obtained from Open Biosystems. A CBF-β complementary DNA matching NM_001755.2 was cloned from the CEM T cell line by RT–PCR. Immunoblots used antibodies to A3G (National Institutes of Health (NIH) ARRRP 10201 courtesy of J. Lingappa), CBF-β (Santa Cruz Biotechnology), HA (HA.11; Covance), TUB (tubulin; Covance), c-Myc (Sigma), Vif (NIH ARRRP 2221 courtesy of D. Gabuzda) and p24/capsid (NIH ARRRP 3537 courtesy of B. Chesebro and K. Wehrly). Details are provided in the Supplementary Methods.

**Received 26 March; accepted 1 November 2011.**

**Published online 21 December 2011.**

1. Goff, S. P. Retrovirus restriction factors. *Mol. Cell* **16,** 849–859 (2004).
2. Malim, M. H. & Emerman, M. HIV-1 accessory proteins—ensuring viral survival in a hostile environment. *Cell Host Microbe* **3,** 388–398 (2008).
3. Albin, J. S. & Harris, R. S. Interactions of host APOBEC3 restriction factors with HIV-1 *in vivo*: implications for therapeutics. *Expert Rev. Mol. Med.* **12,** e4 (2010).
4. Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302,** 1056–1060 (2003).
5. Mehle, A., Goncalves, J., Santa-Marta, M., McPike, M. & Gabuzda, D. Phosphorylation of a novel SOCS-box regulates assembly of the HIV-1 Vif-Cul5 complex that promotes APOBEC3G degradation. *Genes Dev.* **18,** 2861–2866 (2004).
6. Mehle, A. *et al.* Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J. Biol. Chem.* **279,** 7792–7798 (2004).
7. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418,** 646–650 (2002).
8. Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Med.* **9,** 1404–1407 (2003).
9. Conticello, S. G., Harris, R. S. & Neuberger, M. S. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr. Biol.* **13,** 2009–2013 (2003).
10. Stopak, K., de Noronha, C., Yonemoto, W. & Greene, W. C. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **12,** 591–601 (2003).
11. Passmore, L. A. & Barford, D. Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochem. J.* **379,** 513–525 (2004).
12. Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19,** 94–102 (2000).
13. Stanley, B. J. *et al.* Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *J. Virol.* **82,** 8656–8663 (2008).
14. Jäger, S. *et al.* Purification and characterization of HIV-human protein complexes. *Methods* **53,** 13–19 (2011).
15. Jäger, S. *et al.* Global landscape of HIV–human protein complexes. *Nature* doi:10.1038/nature10719 (this issue).
16. Kamura, T. *et al.* VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev.* **18,** 3055–3065 (2004).
17. Huang, D. T. *et al.* E2-RING expansion of the NEDD8 cascade confers specificity to cullin modification. *Mol. Cell* **33,** 483–495 (2009).
18. Wong, W. F., Kohu, K., Chiba, T., Sato, T. & Satake, M. Interplay of transcription factors in T-cell differentiation and function: the role of Runx. *Immunology* **132,** 157–164 (2011).
19. Petroski, M. D. & Deshaies, R. J. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* **123,** 1107–1120 (2005).
20. Kim, H. T. *et al.* Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol. Chem.* **282,** 17375–17386 (2007).
21. DeHart, J. L., Bosque, A., Harris, R. S. & Planelles, V. Human immunodeficiency virus type 1 Vif induces cell cycle delay via recruitment of the same E3 ubiquitin ligase complex that targets APOBEC3 proteins for degradation. *J. Virol.* **82,** 9265–9272 (2008).
22. Stenglein, M. D., Burns, M. B., Li, M., Lengyel, J. & Harris, R. S. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nature Struct. Mol. Biol.* **17,** 222–229 (2010).
23. Hultquist, J. F. *et al.* Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J. Virol.* **85,** 11220–11234 (2011).
24. Rodrigo-Brenni, M. C. & Morgan, D. O. Sequential E2s drive polyubiquitin chain assembly on APC targets. *Cell* **130,** 127–139 (2007).
25. Wu, K., Kovacev, J. & Pan, Z. Q. Priming and extending: a UbcH5/Cdc34 E2 handoff mechanism for polyubiquitination on a SCF substrate. *Mol. Cell* **37,** 784–796 (2010).

26. Koning, F. A. *et al.* Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.* **83,** 9474–9485 (2009).
27. Refsland, E. W. *et al.* Quantitative profiling of the full *APOBEC3* mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.* **38,** 4274–4284 (2010).
28. Saha, A. & Deshaies, R. J. Multimodal activation of the ubiquitin ligase SCF by Nedd8 conjugation. *Mol. Cell* **32,** 21–31 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank members of the Krogan, Gross and Harris laboratories for comments, J. R. Johnson for mass spectrometry, B. Leonard for sharing unpublished data, M. Shales for help with figures, and B. Chesebro, D. Gabuzda, J. Lingappa, M. Malim, K. Wehrly, X. F. Yu, A. Bullock, B. Schulman and the AIDS Research and Reference Reagent Program for reagents. This research was funded by grants from QB3 at University of California, San Francisco, and the National Institutes of Health (P50 GM082250, P01 AI090935 and P50 GM081879 to N.J.K.; U54 RR022220 to A.S.; R01 AI064046 and P01 GM091743 to R.S.H.; P50 GM082250 to J.D.G. and C.S.C.; P41RR001614 and P50GM081879 to A.B.). N.J.K. is a Searle Scholar and a Keck Young Investigator.

**Author Contributions** K.S., R.S.L. and E.K. made equal secondary contributions to this work. S.J. and K.F.S. generated the Vif protein–protein interaction map (Fig. 1); S.J. performed the co-immunoprecipitation confirmation assays (Supplementary Fig. 1); P.C. developed and implemented the MiST scoring system (Fig. 1c); S.J. and E.K. performed double purification analyses (Fig. 1c, d); C.M. and J.K. performed immunoprecipitation analyses (Fig. 1f, g); D.Y.K., L.Y. and D.S. reconstituted the Vif E3 ligase from recombinant components and performed ubiquitination assays (Fig. 2 and Supplementary Figs 2–4); M.L. expressed and purified A3 proteins and did *in vitro* pulldowns (Fig. 2); K.S., J.F.H. and R.S.L. performed CBF-β knockdown, complementation and virus infectivity experiments (Fig. 3 and Supplementary Figs 5–7); and B.D.A. and J.F.H. did the CBF-β–Vif co-immunoprecipitation experiments (Supplementary Fig. 8). A.B., A.S., C.S.C., R.S.H., J.D.G. and N.J.K. supervised the research; S.J., D.Y.K., J.F.H., K.S., R.S.H., J.D.G. and N.J.K. wrote and revised the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.S.H. (rsh@umn.edu), J.D.G. (jdgross@cgl.ucsf.edu) or N.J.K. (krogan@cmp.ucsf.edu).

doi:10.1038/nature10718

# T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction

Wenyan Zhang[1]*, Juan Du[1,2]*, Sean L. Evans[2], Yunkai Yu[2] & Xiao-Fang Yu[1,2]

**The human APOBEC3 cytidine deaminases are potent inhibitors of diverse retroviruses, including human immunodeficiency virus-1 (HIV-1)[1–6]. HIV-1 Vif forms an E3 ubiquitin ligase complex with cullin 5 (CUL5), elongin B and elongin C [7–9], which promotes the polyubiquitination and degradation of APOBEC3 substrates[7,10–14]. Here we demonstrate in human T cells that core binding factor β (CBF-β) is a key regulator of the evasion of HIV-1 from the host defence mediated by APOBEC3. CBF-β, the non-DNA-binding subunit of a heterodimeric transcription factor, regulates the folding and DNA-binding activity of partner RUNX family proteins, which have important roles in the development and differentiation of diverse cell types, including T lymphocytes[15,16]. In our study, knockdown of endogenous CBF-β blocked Vif-induced APOBEC3G polyubiquitination and degradation. CBF-β was not required for the interaction between Vif and APOBEC3G, yet was essential for the assembly of the Vif–CUL5 E3-ubiquitin-ligase complex. CBF-β proved to be a unique regulator of primate lentiviral Vif and not a general component of the CUL5 E3 ubiquitin ligase. We show that Vif and CBF-β physically interact, and that the amino-terminal region of Vif is required for this interaction. Furthermore, interactions with Vif required regions in CBF-β that are not involved in RUNX protein binding[17–19]. Considering the importance of the interaction between Vif and CBF-β, disrupting this interaction represents an attractive pharmacological intervention against HIV-1.**

To identify novel cellular proteins involved in HIV-1 Vif function, we infected human $CD4^+$ T cells (H9) with HIV-1 Vif–haemagglutinin (HA) or HIV-1 (control) virus[7] and then characterized the interaction of cellular proteins with HIV-1 Vif–HA in the infected cells by co-immunoprecipitation analysis. HIV-1 Vif forms an E3 ubiquitin ligase with CUL5 and elongin B/elongin C (ELOB/C) to promote the polyubiquitination and degradation of various substrates, including APOBEC3 and Vif itself[8,20]. To limit the potential degradation of these co-factors, we treated the cells with the proteasome inhibitor MG132 before co-immunoprecipitation. HA-tagged Vif was immunoprecipitated with anti-HA antibody from HIV-1 Vif–HA-infected H9 cell lysates but not HIV-1 (control)-infected H9 cell lysates (Fig. 1a). Immunoprecipitation of Vif–HA was confirmed by immunoblotting with a HA-specific antibody (Fig. 1b). Untagged Vif was not immunoprecipitated from control-infected H9 cells (Fig. 1b, lane 2). CUL5 (84 kDa) was co-precipitated with Vif–HA (Fig. 1a, lane 1) and confirmed by immunoblotting with a CUL5-specific antibody (Fig. 1b, lane 1); ELOB/C was also co-precipitated with Vif–HA (Fig. 1b, lane 1). These proteins were not co-precipitated from control HIV-1-infected H9 cells (Fig. 1b, lane 2).

A previously uncharacterized protein of ~22 kDa also co-precipitated with Vif–HA (Fig. 1a, lane 1). This protein was identified by mass spectrometry analysis as CBF-β, a ubiquitously expressed host protein[15,16]. The identification of the protein was confirmed by immunoblotting with a CBF-β-specific antibody (Fig. 1b, lane 1). We also detected an interaction between HIV-1 Vif–HA and CBF-β in transfected 293T (human embryonic kidney) cells (Fig. 1c, lane 4), indicating that the interaction between CBF-β and HIV-1 Vif can occur in the absence of

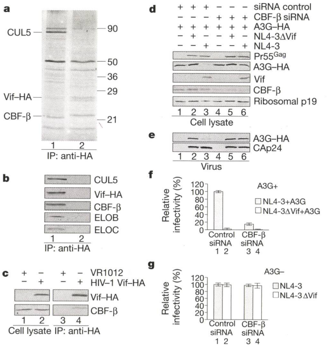


**Figure 1 | Characterization of the functional interaction between HIV-1 Vif and CBF-β. a,** Co-immunoprecipitation of cellular proteins with Vif–HA. Cell lysates from H9 cells infected with HIV-1 Vif–HA (HXB2Neo-Vif-HA, lane 1) or with control HIV-1 (that is, HXB2Neo containing untagged Vif, lane 2) were immunoprecipitated (IP) with anti-HA antibody conjugated to agarose beads, followed by SDS–PAGE and silver staining. The identification of CBF-β co-precipitated with Vif–HA (lane 1) was achieved by mass spectrometry. **b,** Immunoblot of precipitated samples from HIV-1- or HIV-1 Vif–HA-infected H9 cells using antibodies against CUL5, CBF-β, ELOB, ELOC or HA (to detect Vif–HA). **c,** Co-immunoprecipitation of endogenous CBF-β with Vif–HA from transfected 293T-cell samples. **d,** Vif-mediated downregulation of APOBEC3G (A3G) requires CBF-β. 293T cells were co-transfected with expression vectors for APOBEC3G–HA plus NL4-3, NL4-3ΔVif, or a control vector (pcDNA3.1) in the presence of control siRNA (lanes 1–3) or siRNA against CBF-β (lanes 4–6). **e,** CBF-β is required for Vif-mediated exclusion of APOBEC3G from virions. **e, f,** Supernatants from transfected 293T cells as described in **d** were collected and analysed for APOBEC3G–HA virion packaging (**e**) and viral infectivity (**f**). **f,** Virus infectivity was assessed by multinuclear activation of a galactosidase indicator (MAGI) assay. Virus input was normalized by the amount of CAp24. Virus infectivity with NL4-3 (Vif-positive) produced from transfected 293T cells in the presence of control siRNA was set to 100%. Error bars in **f** and **g** represent the standard deviations from triplicate wells. **g,** Silencing CBF-β expression had no effect on viral infectivity in the absence of APOBEC3G.

---

[1]First Hospital of Jilin University, Institute of Virology and AIDS Research, Changchun, Jilin Province 130021, China. [2]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, Maryland 21205, USA.
*These authors contributed equally to this work.

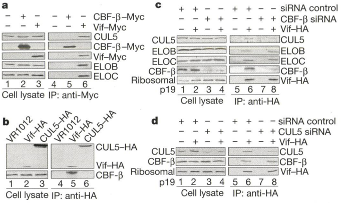

a
CBF-β–Myc
Vif–Myc
CUL5
CBF-β–Myc
Vif–Myc
ELOB
ELOC
1 2 3 4 5 6
Cell lysate IP: anti-Myc
b
VR1012
Vif–HA
CUL5–HA
VR1012
Vif–HA
CUL5–HA
CUL5–HA
Vif–HA
CBF-β
1 2 3 4 5 6
Cell lysate IP: anti-HA
c
siRNA control
CBF-β siRNA
Vif–HA
CUL5
ELOB
ELOC
CBF-β
Ribosomal p19
CUL5
ELOB
ELOC
CBF-β
Vif–HA
1 2 3 4 5 6 7 8
Cell lysate IP: anti-HA
d
siRNA control
CUL5 siRNA
Vif–HA
CUL5
CBF-β
Ribosomal p19
CUL5
CBF-β
Vif–HA
1 2 3 4 5 6 7 8
Cell lysate IP: anti-HA

**Figure 2 | CBF-β is critical for the interaction between Vif and CUL5.** **a–d**, 293T cells were transfected with the indicated expression vectors. Cell lysates were prepared and immunoprecipitated followed by immunoblot analysis as described in Fig. 1. **a**, Co- immunoprecipitation of CUL5, ELOB and ELOC with Vif–Myc but not CBF-β–Myc. **b**, Co-immunoprecipitation of endogenous CBF-β with Vif–HA but not CUL5–HA. **c**, Silencing CBF-β inhibits the interaction between Vif and CUL5. **d**, Silencing CUL5 does not affect the Vif–CBF-β interaction.

other viral proteins. MG132 treatment increased the stability of HIV-1 Vif but did not alter its interaction with CBF-β (Supplementary Fig. 1); RNase treatment also did not block its interaction with CBF-β. Using a live cell imaging approach, we observed co-localization of yellow fluorescent protein (YFP)–Vif and cyan fluorescent protein (CFP)–CBF-β in transfected 293T cells along with a positive fluorescence resonance energy transfer (FRET) signal, further validating a physical interaction (Supplementary Fig. 2).

To determine whether CBF-β is required for the function of HIV-1 Vif, we knocked down endogenous CBF-β expression by CBF-β-specific short interfering RNA (siRNA). As compared to control siRNA-treated cells (Fig. 1d, lanes 1–3), CBF-β was reduced by ~70% in 293T cells treated with CBF-β-specific siRNA (lanes 4–6). In the control siRNA-treated cells, HIV-1 Vif efficiently reduced the expression of APOBEC3G (lane 3) when compared to HIV-1 NL4-3 strain Vif-negative (NL4-3ΔVif) virus-expressing cells (lane 2). Silencing of endogenous CBF-β (lanes 4–6) impeded the ability of Vif to reduce the expression of APOBEC3G (lane 6). Because HIV-1 Vif mainly reduces the expression of APOBEC3G by inducing its degradation, these data indicated that CBF-β is required for the HIV-1 Vif-mediated degradation of APOBEC3G (Fig. 1d, compare lanes 2 and 3 to lanes 5 and 6). Interestingly, expression of exogenous CFP–CBF-β enhanced both wild-type, untagged Vif (pcDNA-hVif)- and YFP–Vif-induced APOBEC3G degradation in 293T cells (Supplementary Fig. 2).

To exert its antiviral activity, APOBEC3G must be packaged into HIV-1 virions so that it can then inhibit virus infectivity in new target cells; Vif suppresses this antiviral activity by limiting the packaging of APOBEC3G into virions. Indeed, we found that when Vif was present, APOBEC3G was efficiently excluded from released HIV-1 virions in control siRNA-treated cells (Fig. 1e, lane 3); when Vif was absent, APOBEC3G could be packaged into HIV-1 virions (Fig. 1e, lane 2). However, the ability of Vif to exclude APOBEC3G from released HIV-1 virions was markedly compromised when CBF-β expression was knocked down in virus-producing cells (Fig. 1e, compare lanes 3 and 6). Reducing the expression of CBF-β specifically blocked the ability of HIV-1 Vif to suppress the antiviral activity of APOBEC3G (Fig. 1f). In contrast, silencing CBF-β expression had no effect on the viral infectivity of NL4-3 or NL4-3ΔVif when APOBEC3G was absent (Fig. 1g), indicating that CBF-β regulates HIV-1 infectivity only in the presence of APOBEC3G.

CBF-β apparently does not interact directly with APOBEC3G and is not required for the Vif–APOBEC3G interaction (Supplementary Fig. 3). Human APOBEC3F is another human cytidine deaminase that has potent anti-HIV-1 activity; however, APOBEC3F and APOBEC3G are recognized by HIV-1 Vif through distinct mechanisms[21–27]. We postulated that if CBF-β is not involved in Vif-substrate recognition, it most likely regulates Vif mediated suppression of APOBEC3F as well as APOBEC3G. Indeed, HIV-1 Vif-mediated degradation, virion exclusion and suppression of APOBEC3F also required CBF-β (Supplementary Fig. 4). Thus, CBF-β is apparently required for Vif activity at a critical step(s) other than substrate recognition.

Because CBF-β does not affect the interaction between HIV-1 Vif and APOBEC3G, it is possible that CBF-β influences the ubiquitin ligase activity of Vif–CUL5 E3 and could in fact be a previously unrecognized component of the CUL5–ELOB/C E3 ubiquitin ligase. To determine whether CBF-β is an integral component of this complex, we examined the interaction of CBF-β with CUL5 and ELOB/C. Co-immunoprecipitation of CUL5 and ELOB/C with HIV-1 Vif–Myc (Fig. 2a, lane 6) was readily detected, but there was no co-immunoprecipitation of CUL5 or ELOB/C with CBF-β–Myc under the same conditions (lane 5). We also saw no interaction between endogenous CBF-β and CUL5–HA in a reciprocal co-immunoprecipitation analysis (Fig. 2b, lane 6), despite a readily detectable interaction between CBF-β and Vif–HA (lane 5). Therefore, CBF-β is unlikely to be an integral component of the CUL5–ELOB/C E3 complex under physiological conditions.

To determine whether CBF-β might have a role in the assembly of the Vif–CUL5 E3 ubiquitin ligase, we examined the interaction of Vif–HA with endogenous CUL5 and ELOB/C by co-immunoprecipitation analysis. Reducing CBF-β expression with specific siRNA had no effect on the intracellular expression of CUL5 or ELOB/C (Fig. 2c, compare lanes 1 and 2 to lanes 3 and 4). CUL5 and ELOB/C were co-immunoprecipitated with HIV-1 Vif–HA by the anti-HA antibody in control siRNA-treated cells (lane 6), but not in control cells that did not express HIV-1 Vif–HA (lanes 5 and 7). Co-immunoprecipitation of CUL5 with HIV-1 Vif–HA was significantly inhibited when CBF-β was silenced (Fig. 2c, lane 8) but co-precipitation of ELOB/C with HIV-1 Vif–HA was not (lane 8). Although silencing of CBF-β inhibited the interaction of Vif with CUL5 (Fig. 2c), reciprocal silencing of CUL5 (Fig. 2d, lane 4) had no

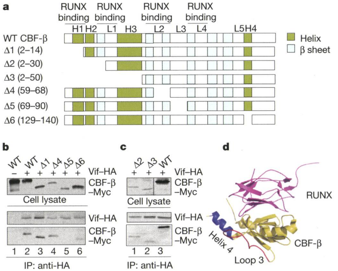

a
RUNX binding
RUNX binding
RUNX binding
RUNX binding
H1 H2 L1 H3 L2 L3 L4 L5 H4
WT CBF-β
Δ1 (2–14)
Δ2 (2–30)
Δ3 (2–50)
Δ4 (59–68)
Δ5 (69–90)
Δ6 (129–140)
🟩 Helix
☐ β sheet
b
WT
WT
Δ1
Δ4
Δ5
Δ6
Vif–HA
− + + + + +
CBF-β –Myc
Cell lysate
Vif–HA
CBF-β –Myc
1 2 3 4 5 6
IP: anti-HA
c
Δ2
Δ3
WT
Vif–HA
+ + +
CBF-β –Myc
Cell lysate
Vif–HA
CBF-β –Myc
1 2 3
IP: anti-HA
d
RUNX
Helix 4
CBF-β
Loop 3

**Figure 3 | Identification of regions in CBF-β is important for HIV-1 Vif interaction.** **a**, Diagrams of mutant CBF-β constructs. The numbers in parentheses indicate amino acids that were deleted in the mutant CBF-β constructs. **b**, **c**, Co-precipitation of full-length (wild-type; WT) CBF-β–Myc and CBF-β mutants with HIV-1 Vif–HA. Cell lysates and precipitated samples were analysed by immunoblotting with an anti-Myc or anti-HA antibody. **d**, The structures of CBF-β (yellow) complexed with RUNX1 (pink) have been well characterized[17–19]. Loop 3 (red) and helix 4 (blue) of CBF-β, which are located away from the interfaces between CBF-β and RUNX1, are important for the Vif–CBF-β interaction.

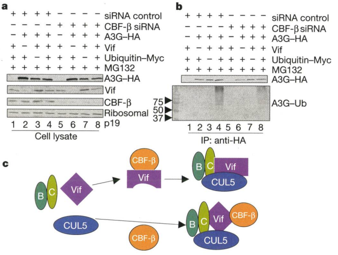


**Figure 4 | Vif-induced polyubiquitination of APOBEC3G requires CBF-β.** **a–c**, 293T cells were transfected with an expression vector encoding APOBEC3G–HA (A3G–HA), Vif, Myc-tagged ubiquitin (Ub) or control vector as indicated. All experiments were performed in the presence of MG132. **a**, Cells were lysed and analysed by immunoblot analysis using an anti-Vif, anti-CBF-β or anti-HA antibody. **b**, Cell lysates were immunoprecipitated with the anti-HA antibody and analysed by immunoblotting with an anti-Myc antibody to detect ubiquitinated APOBEC3G. Unmodified APOBEC3G–HA was detected with the anti-HA antibody. **c**, Proposed models for the role of CBF-β in HIV-1 Vif function.

detectable effect on the CBF-β–Vif interaction (Fig. 2d, lane 8). Again, these data indicate that CBF-β is not a general co-factor of CUL5, and CBF-β uniquely regulates the interaction of HIV-1 Vif with CUL5.

We also observed that the activity of the adenoviral E4orf6-containing CUL5–ELOB/C ubiquitin ligase was not affected by CBF-β silencing (Supplementary Fig. 5). In addition, cellular SOCS-box-containing proteins (that is, ASB6 and SOCS3) (Supplementary Fig. 5), which are known to assemble with the CUL5–ELOB/C E3 ubiquitin ligase, did not interact with CBF-β. Also, although primate lentiviral Vif interacted with CBF-β, Vif from bovine immunodeficiency virus (BIV) did not (Supplementary Fig. 6). BIV Vif-mediated bovine APOBEC3 depletion was not affected by CBF-β silencing.

CBF-β is a ubiquitously expressed component of a heterodimeric transcription factor including RUNX family proteins[15,16]. The interface between CBF-β and RUNX proteins is well defined[17–19]. To explore the regions of CBF-β and HIV-1 Vif that are critical for their interaction, we first examined the Vif-binding activity of a series of in-frame CBF-β deletion mutants (Fig. 3a). The regions of CBF-β that were important for the interaction with HIV-1 Vif mapped to loop 3 and helix 4 (Fig. 3b, c), regions that do not interact with RUNX proteins (Fig. 3d).

Consistent with the idea that HIV-1 Vif does not compete with RUNX for binding to CBF-β, we observed that the interaction of RUNX1 with CBF-β was not affected by the presence of Vif (Supplementary Fig. 7); in fact, a possible co-precipitation of Vif–HA with RUNX1–Myc was detected (Supplementary Fig. 7). RUNX1-mediated transcriptional activity[28] was also not significantly affected by Vif, as measured in a CBF promoter-driven luciferase reporter assay (Supplementary Fig. 7). We determined that the N-terminal regions of Vif, including amino acids W21 and W38, were important for CBF-β binding (Supplementary Fig. 8).

To determine whether CBF-β can influence the HIV-1 Vif-induced polyubiquitination of APOBEC3G, we transfected 293T cells with various combinations of expression vectors for HIV-1 Vif–Myc, APOBEC3G–HA and Myc-tagged ubiquitin in the presence of control siRNA or siRNA against CBF-β (Fig. 4a). To inhibit APOBEC3G degradation, all transfected cells were treated with MG132. HA-tagged APOBEC3G was immunoprecipitated by the anti-HA antibody, followed by the detection of ubiquitinated APOBEC3G–HA using the antibody against Myc-tagged ubiquitin. Polyubiquitinated APOBEC3G–HA was detected in cells expressing Vif (Fig. 4b, lane 4); however, the Vif-induced polyubiquitination of APOBEC3G was inhibited by siRNA against CBF-β (lane 8).

Taken together, our results demonstrate that CBF-β is a key regulator of HIV-1 Vif function. We have further shown that CBF-β is critical for Vif-induced endogenous APOBEC3G degradation and HIV-1 replication in H9 CD4$^+$ T cells (Supplementary Fig. 9). CBF-β allosterically regulates the conformation and DNA-binding activity of RUNX proteins[17–19]. It remains to be determined, however, whether CBF-β can induce conformational changes in HIV-1 Vif that allow CUL5 binding to occur. Alternatively, CBF-β may stabilize the interaction between Vif and CUL5 in the E3 ubiquitin ligase complex (Fig. 4c). Vif–CBF-β binding is critical for HIV-1 infection but does not seem to interfere with normal CBF-β–RUNX interaction. Thus, the interface between Vif and CBF-β may constitute a promising target for new anti-HIV drug discovery.

## METHODS SUMMARY

**Live-cell imaging and FRET assay.** 293T cells were transfected with YFP–Vif (2 μg) and CFP–CBF-β (0.25 μg) using Lipofectamine 2000 (Invitrogen) in 6-well coverslip glass-bottom cell culture dishes (MatTek) at 37 °C with 5% $CO_2$. Eightfold excess YFP–Vif was used due to low expression of YFP–Vif compared to CFP–CBF-β (Supplementary Fig. 2d). YFP–CFP dimer (0.25 μg), co-transfection of YFP (0.25 μg) plus CFP (0.25 μg), and CFP–CBF-β alone were used for a positive and negative FRET signal as imaging controls. Cells were visualized 24–36 h after transfection using a Zeiss LSM510-Meta confocal imaging system equipped with four argon lasers (458, 477, 488 and 514 nm lines), two HeNe lasers (542 and 633 nm), and one diode laser (405 nm). YFP and CFP were excited using a 514 nm and 458 nm laser, respectively. After excitation, channel mode images were collected using a long-pass filter for CFP emission (LP 420) or band-pass filter for YFP emission (BP 530-600). In addition, the channel mode detector gain was set so that neither CFP nor YFP images contained saturated pixels before bleaching. A total of ten images (five pre-bleach and five post-bleach) were acquired for each cell with a ×63 objective with ×2 zoom, and image analysis and manipulation was performed using Zeiss LSM imaging software. FRET analysis was performed using the acceptor photo-bleaching technique: YFP or YFP–Vif was photo-bleached for approximately 5 min at 514 nm (100% laser power), and the average donor signal was evaluated before and after acceptor photobleaching. As a control, cells were also imaged under similar conditions, except the 514 nm laser was set at 0% transmission to prevent acceptor photobleaching. Two experiments were performed for each bleached sample for a total of ten cells each. One experiment was performed for each unbleached sample for a total of four cells each. The FRET efficiency percentage ($E$) was calculated using the FRET macro of Zeiss LSM software. The following equation was used: $E = (F_{D\Delta}/F_{Dpost}) \times 100$, where $F_{D\Delta}$ is the change in donor (CFP) fluorescence intensity before and after acceptor (YFP) bleaching and $F_{Dpost}$ is donor fluorescence intensity after acceptor bleaching. The FRET macro of Zeiss LSM software also subtracts background when calculating $E$.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 27 May; accepted 17 November 2011.**

**Published online 21 December 2011.**

1. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418,** 646–650 (2002).
2. Navarro, F. & Landau, N. R. Recent insights into HIV-1 Vif. *Curr. Opin. Immunol.* **16,** 477–482 (2004).
3. Cullen, B. R. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J. Virol.* **80,** 1067–1076 (2006).
4. Chiu, Y. L. & Greene, W. C. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.* **26,** 317–353 (2008).
5. Malim, M. H. & Emerman, M. HIV-1 accessory proteins—ensuring viral survival in a hostile environment. *Cell Host Microbe* **3,** 388–398 (2008).
6. Goila-Gaur, R. & Strebel, K. HIV-1 Vif, APOBEC, and intrinsic immunity. *Retrovirology* **5,** 51 (2008).
7. Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-CUL5-SCF complex. *Science* **302,** 1056–1060 (2003).

8. Mehle, A., Goncalves, J., Santa-Marta, M., McPike, M. & Gabuzda, D. Phosphorylation of a novel SOCS-box regulates assembly of the HIV-1 Vif-CUL5 complex that promotes APOBEC3G degradation. *Genes Dev.* **18,** 2861–2866 (2004).
9. Yu, Y., Xiao, Z., Ehrlich, E. S., Yu, X. & Yu, X. F. Selective assembly of HIV-1 Vif-CUL5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev.* **18,** 2867–2872 (2004).
10. Stopak, K., de Noronha, C., Yonemoto, W. & Greene, W. C. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **12,** 591–601 (2003).
11. Marin, M., Rose, K. M., Kozak, S. L. & Kabat, D. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature Med.* **9,** 1398–1403 (2003).
12. Conticello, S. G., Harris, R. S. & Neuberger, M. S. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr. Biol.* **13,** 2009–2013 (2003).
13. Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Med.* **9,** 1404–1407 (2003).
14. Mehle, A. *et al.* Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J. Biol. Chem.* **279,** 7792–7798 (2003).
15. de Bruijn, M. F. & Speck, N. A. Core-binding factors in hematopoiesis and immune function. *Oncogene* **23,** 4238–4248 (2004).
16. Ito, Y. RUNX genes in development and cancer: regulation of viral gene expression and the discovery of RUNX family genes. *Adv. Cancer Res.* **99,** 33–76 (2008).
17. Tahirov, T. H. *et al.* Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFβ. *Cell* **104,** 755–767 (2001).
18. Bravo, J., Li, Z., Speck, N. A. & Warren, A. J. The leukemia-associated AML1 (Runx1)–CBFβ complex functions as a DNA-induced molecular clamp. *Nature Struct. Biol.* **8,** 371–378 (2001).
19. Yan, J., Liu, Y., Lukasik, S. M., Speck, N. A. & Bushweller, J. H. CBFβ allosterically regulates the Runx1 Runt domain via a dynamic conformational equilibrium. *Nature Struct. Mol. Biol.* **11,** 901–906 (2004).
20. Liu, B., Sarkis, P. T., Luo, K., Yu, Y. & Yu, X. F. Regulation of Apobec3F and human immunodeficiency virus type 1 Vif by Vif-CUL5-ElonB/C E3 ubiquitin ligase. *J. Virol.* **79,** 9579–9587 (2005).
21. Simon, V. *et al.* Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog.* **1,** e6 (2005).
22. Tian, C. *et al.* Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F. *J. Virol.* **80,** 3112–3115 (2006).
23. Schröfelbauer, B., Senger, T., Manning, G. & Landau, N. R. Mutational alteration of human immunodeficiency virus type 1 Vif allows for functional interaction with nonhuman primate APOBEC3G. *J. Virol.* **80,** 5984–5991 (2006).
24. Russell, R. A. & Pathak, V. K. Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. *J. Virol.* **81,** 8201–8210 (2007).
25. Mehle, A. *et al.* Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *J. Virol.* **81,** 13235–13241 (2007).
26. He, Z., Zhang, W., Chen, G., Xu, R. & Yu, X. F. Characterization of conserved motifs in HIV-1 Vif required for APOBEC3G and APOBEC3F interaction. *J. Mol. Biol.* **381,** 1000–1011 (2008).
27. Zhang, W., Chen, G., Niewiadomska, A. M., Xu, R. & Yu, X. F. Distinct determinants in HIV-1 Vif and human APOBEC3 proteins are required for the suppression of diverse host anti-viral proteins. *PLoS ONE* **3,** e3963 (2008).
28. Cao, W., Adya, N., Britos-Bray, M., Liu, P. P. & Friedman, A. D. The core binding factor (CBF) α interaction domain and the smooth muscle myosin heavy chain (SMMHC) segment of CBFβ-SMMHC are both required to slow cell proliferation. *J. Biol. Chem.* **273,** 31534–31540 (1998).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank K. Strebel, A. Friedman, N. Speck, D. Yue, R. Siliciano, M. Malim, R. Harris, T. Inoue and D. Gabuzda for critical reagents; T. Wang, K. Zhao, X. Zhou and A. Zhen for technical assistance; R. Markham, J. Margolick and J. Bream for thoughtful discussions; and D. McClellan for editorial assistance. We also wish to thank staff within the Mass Spectrometry Core and the Institute for Basic Biomedical Sciences Microscope Facility at Johns Hopkins School of Medicine for their technical assistance. The following reagents were obtained through the AIDS Research and Reference Reagents Program, Division of AIDS, NIAID, NIH: monoclonal antibodies against HIV-1 p24 (B. Chesebro and H. Chen), pNL4-3 (M. Martin), pcDNA-hVif (S. Bour and K. Strebel), antiserum to HIV-1 Vif (D. Gabuzda) and MAGI-CCR5 cells (J. Overbaugh). This work was supported in part by funding from the Chinese Ministry of Science and Technology (2012CB911100) and Chinese Ministry of Education (IRT1016), the Key Laboratory of Molecular Virology, Jilin Province (20102209), China, and a grant (2R56AI62644-6) from the NIAID.

**Author Contributions** W.Z., J.D., S.L.E. and Y.Y. performed experiments and analysed the data. X.-F.Y. directed the project, analysed the data, and wrote the paper with help from all authors.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to X-F.Y. (xfyu@jhsph.edu).

doi:10.1038/nature10602

# Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia

Christian M. Metallo[1]†, Paulo A. Gameiro[1,2,3,4], Eric L. Bell[5], Katherine R. Mattaini[5,6], Juanjuan Yang[3,4], Karsten Hiller[1]†, Christopher M. Jewell[6], Zachary R. Johnson[6], Darrell J. Irvine[6,7], Leonard Guarente[5], Joanne K. Kelleher[1], Matthew G. Vander Heiden[5,6,8], Othon Iliopoulos[3,4] & Gregory Stephanopoulos[1]

**Acetyl coenzyme A (AcCoA) is the central biosynthetic precursor for fatty-acid synthesis and protein acetylation. In the conventional view of mammalian cell metabolism, AcCoA is primarily generated from glucose-derived pyruvate through the citrate shuttle and ATP citrate lyase in the cytosol[1–3]. However, proliferating cells that exhibit aerobic glycolysis and those exposed to hypoxia convert glucose to lactate at near-stoichiometric levels, directing glucose carbon away from the tricarboxylic acid cycle and fatty-acid synthesis[4]. Although glutamine is consumed at levels exceeding that required for nitrogen biosynthesis[5], the regulation and use of glutamine metabolism in hypoxic cells is not well understood. Here we show that human cells use reductive metabolism of α-ketoglutarate to synthesize AcCoA for lipid synthesis. This isocitrate dehydrogenase-1 (IDH1)-dependent pathway is active in most cell lines under normal culture conditions, but cells grown under hypoxia rely almost exclusively on the reductive carboxylation of glutamine-derived α-ketoglutarate for *de novo* lipogenesis. Furthermore, renal cell lines deficient in the von Hippel–Lindau tumour suppressor protein preferentially use reductive glutamine metabolism for lipid biosynthesis even at normal oxygen levels. These results identify a critical role for oxygen in regulating carbon use to produce AcCoA and support lipid synthesis in mammalian cells.**

Although hypoxic cells exhibit a shift towards glycolytic metabolism[4], a functional electron transport chain and glutamine-derived carbon are required for proliferation of most transformed cells[6]. In line with these studies, we observed increased glutamine consumption when A549 cells were cultured at approximately 1% oxygen while glutamate secretion remained unchanged (Supplementary Fig. 1a), indicating that net glutamine consumption was elevated and suggesting that glutamine carbon is used for biosynthesis. Notably, glucose consumption and lactate secretion also increased in these experiments. Consistent with this observation, we found that proliferating cells under both normoxia and hypoxia incorporate glutamine-derived carbon into lipids (Supplementary Fig. 1b). Glutamine can contribute carbon to lipogenic AcCoA through two distinct pathways. Cells can oxidatively metabolize glutamine-derived α-ketoglutarate (αKG) in the tricarboxylic acid (TCA) cycle and generate pyruvate from malate by glutaminolysis[5]. Alternatively, some tissues can reductively carboxylate αKG to generate citrate[7,8], and recent studies have indicated that the IDH reaction is highly reversible[9–11]. To determine which pathway cells use to incorporate glutamine carbon into lipids we used stable isotopic tracers[12–14]. We first cultured several cancer cell lines with [1-$^{13}$C]glutamine under normoxia and quantified the isotopic label present in metabolite pools along this pathway (Fig. 1a, red carbon atoms). All cells tested with this tracer retained significant label from [1-$^{13}$C]glutamine in citrate and metabolites downstream of the irreversible ATP citrate lyase reaction, indicating that the reductive flux contributes to the cytosolic AcCoA pool (Supplementary Fig. 2). Additional evidence for activity along this pathway was obtained with a uniformly $^{13}$C-labelled ([U-$^{13}$C$_5$]) glutamine tracer (Supplementary Fig. 3).

To quantify the specific contributions of oxidative and reductive glutamine metabolism to fatty-acid synthesis, we cultured cells in the presence of tracers for several days and performed isotopomer spectral analysis (ISA; Supplementary Fig. 4)[15]. Use of [5-$^{13}$C]glutamine specifically allows estimation of the flux of glutamine to lipids through reductive carboxylation (Fig. 1a, blue carbon atoms). Virtually all cell lines cultured with this tracer generated labelled fatty acids, metabolizing glutamine reductively in the TCA cycle to supply 10–25% of their lipogenic AcCoA (Fig. 1b and Supplementary Fig. 5). Consistent with these data, we were able to detect 98 ± 5 c.p.m. per $10^6$ cells in hexane extracts of A549 cells cultured with [5-$^{14}$C]glutamine. Next we directly compared the contribution of glutamine to fatty acids through the reductive flux as a fraction of the total, the latter determined by using [U-$^{13}$C$_5$]glutamine. In all cell lines tested, including those derived from lung, mammary, colon and squamous cell carcinoma as well as melanoma, glioblastoma and leukaemia, [5-$^{13}$C$_5$]glutamine labelled most of the AcCoA derived from glutamine (Fig. 1b), highlighting the general use of reductive carboxylation as the primary route through which glutamine, glutamate and αKG carbon are converted to lipids in cultured cells (Fig. 1b and Supplementary Figs 2, 3 and 5). The glutaminolysis pathway is also an important means of glutamine catabolism and can be characterized by quantifying the contribution of glutamine carbon to lactate. Consistent with published reports[5], glutamine-derived $^{13}$C label was also detected in lactate, and the amount of $^{13}$C-labelled lactate produced was highest in glioblastoma-derived cells compared with other cells cultured with [U-$^{13}$C$_5$]glutamine (Supplementary Fig. 6).

Mammalian cells express three IDH enzymes encoded by separate genes: IDH1 (cytosolic, NADP$^+$-dependent), IDH2 (mitochondrial, NADP$^+$-dependent) and the multi-subunit enzyme IDH3 (mitochondrial, NAD$^+$-dependent). IDH3 is allosterically regulated and is thought to operate in the oxidative direction. The NADP$^+$-dependent isozymes are capable of catalysing the reductive reaction; however, the specific enzyme responsible for this flux is not definitively known[16]. As measurements of metabolite pools in subcellular compartments and labelling therein cannot yet be reliably obtained, we used RNA interference to knockdown expression of IDH1 and IDH2 selectively in A549 cells. Using labelling from [1-$^{13}$C]glutamine as a readout, we measured a significant and robust decrease in reductive carboxylation when IDH1 messenger RNA (mRNA) was targeted using short hairpin RNAs (shRNAs) (Fig. 1c, d). These changes were consistent with results using [U-$^{13}$C$_5$]glutamine (Supplementary Fig. 7) and reproduced

---

[1] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2] Department of Life Sciences, University of Coimbra, 3004-517 Coimbra, Portugal. [3] Massachusetts General Hospital Cancer Center, Boston, Massachusetts 02114, USA. [4] Massachusetts General Hospital Center for Cancer Research, Charlestown, Massachusetts 02129, USA. [5] Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [6] Koch Institute for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [7] Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. [8] Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. †Present addresses: Department of Bioengineering, University of California at San Diego, La Jolla, California 92093, USA (C.M.M.); Luxembourg Centre for Systems Biomedicine, University of Luxembourg, L-1511, Luxembourg (K.H.).

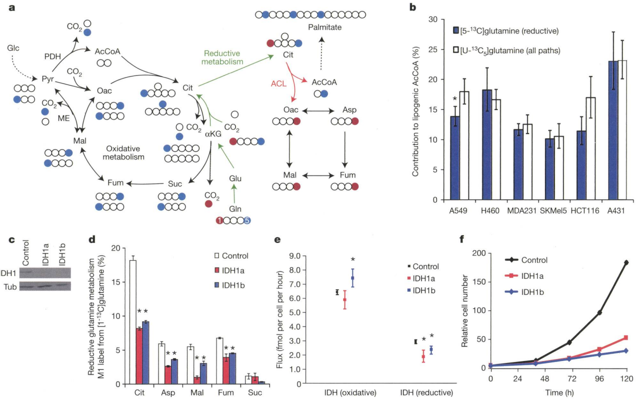


**Figure 1 | Reductive carboxylation is the primary route of glutamine to lipids.** **a**, Schematic of carbon atom (circles) transitions and tracers used to detect reductive glutamine metabolism. Isotopic label from [1-$^{13}$C]glutamine (red) is lost during oxidative conversion to succinate (Suc) but retained on citrate (Cit), oxaloacetate (Oac), aspartate (Asp), malate (Mal) and fumarate (Fum) in the reductive pathway (green arrows). [5-$^{13}$C]glutamine (blue) transfers label to AcCoA through reductive metabolism only. Molecular symmetry is shown for oxidative metabolism. Glc, glucose; Pyr, pyruvate; ME, malic enzyme; PDH, pyruvate dehydrogenase; Glu, glutamate; Gln, glutamine. **b**, Contribution of [5-$^{13}$C]glutamine and [U-$^{13}C_5$]glutamine to lipogenic AcCoA in cell lines. **c**, IDH1 levels in A549 cells expressing IDH1-specific (IDH1a and IDH1b) or control shRNAs. **d**, Metabolite labelling from [1-$^{13}$C]glutamine from cells in **c**. **e**, IDH flux estimates from $^{13}$C metabolic flux analysis model in control or IDH1-knockdown A549 cells cultured with [U-$^{13}C_5$]glutamine. **f**, Cell proliferation of A549 cells expressing control or IDH1-shRNAs. Error bars, 95% confidence intervals (**b**, **e**) and s.e.m. ($n = 3$) (**d**, **f**). $^*P < 0.05$.

using several cell lines (Supplementary Fig. 8). Finally, we used $^{13}$C metabolic flux analysis[17,18] to quantify intracellular fluxes using [U-$^{13}C_5$]glutamine in A549 cells. The fitted data suggested that reductive IDH flux significantly decreased when IDH1 protein levels were decreased, and this change was the primary alteration observed in the network (Fig. 1e; see Supplementary Tables 1–5 for complete results and metabolic flux analysis model description).

Our results suggest that IDH1 can convert NADPH, αKG and $CO_2$ to isocitrate and $NADP^+$ in the cytosol. Enzymatic analysis using recombinant protein indicated that IDH1 is indeed capable of consuming NADPH and is responsive to physiological levels of $CO_2$ (Supplementary Fig. 9). Importantly, the proliferation rate of all cell lines with IDH1 knockdown was impaired (Fig. 1f and Supplementary Fig. 8), indicating that reductive metabolism of αKG in the cytosol may be necessary for robust growth. In contrast to our results with IDH1 shRNAs, we detected no significant change in reductive flux when targeting IDH2 mRNA in A549, MDA-MB-231 and HCT116 cells (Supplementary Fig. 10). Although IDH2 may promote reductive carboxylation in some tissues or conditions, these results are consistent with the proposed role of IDH2 as an oxidative TCA cycle enzyme[19].

Intriguingly, we detected a significant increase in reductive carboxylation activity when culturing cells under hypoxia (Supplementary Fig. 11). As glucose is usually the primary carbon source for mammalian tissues[1–3], we next compared the contributions of reductive glutamine metabolism and glucose oxidation to fatty-acid synthesis by culturing cells with either [5-$^{13}$C]glutamine or uniformly labelled [U-$^{13}C_6$]glucose under normal tissue culture conditions or hypoxia. Strikingly, oxygen levels influenced fatty-acid labelling from both tracers (Fig. 2a, b). Cells preferentially used glucose carbon for palmitate synthesis under normoxic conditions; however, fatty acids produced under hypoxia were primarily synthesized from glutamine carbon through the reductive pathway. In fact, the reductive carboxylation of glutamine-derived αKG accounted for approximately 80% of the carbon used for *de novo* lipogenesis in A549 cells growing under hypoxia (Fig. 2c). Conversely, we detected a concomitant decrease in the contribution of [U-$^{13}C_6$]glucose to fatty-acid synthesis under this condition. Significant increases in the relative use of this pathway were observed in all cell lines tested, including non-transformed cells (Supplementary Fig. 12). In addition, T lymphocytes freshly isolated from a mouse spleen preferentially used reductive glutamine metabolism over glucose oxidation for fatty-acid synthesis when activated under hypoxia (Supplementary Fig. 13). Although proliferation rates and relative *de novo* lipogenesis were lower under hypoxia (Supplementary Fig. 14a, b), the net flux of reductive glutamine metabolism to palmitate synthesis was significantly increased in hypoxic cultures (Fig. 2d). Knockdown of IDH1 protein mitigated the use of reductive glutamine metabolism for lipogenesis under hypoxia (Supplementary Fig. 14c). Although most human cells require glutamine for nucleotide and hexosamine biosynthesis, some cell lines can grow in the absence of exogenous sources of glutamine[20]. Remarkably, we found that hypoxia increases the dependence of such cells on glutamine, as evidenced by decreased proliferation in the absence of glutamine and increased reductive glutamine metabolism under hypoxia when glutamine is present (Fig. 2e and Supplementary Fig. 15).

To gain insight into the mechanisms controlling this switch to reductive glutamine metabolism, we analysed changes in the labelling

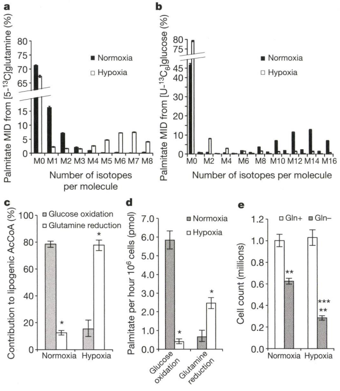


**Figure 2 | Hypoxia reprograms cells to rely on reductive glutamine metabolism for lipid synthesis. a, b,** Labelling of palmitate extracts from A549 cells cultured under normoxia or hypoxia with [5-$^{13}$C]glutamine (**a**) or [U-$^{13}$C$_6$]glucose (**b**). Similar results were observed in myristate, oleate and stearate pools (not shown). **c,** Relative contribution of glucose oxidation ([U-$^{13}$C$_6$]glucose) or glutamine reduction ([5-$^{13}$C]glutamine) to lipogenic AcCoA in A549 cells under normoxia and hypoxia. **d,** Absolute fluxes of [U-$^{13}$C$_6$]glucose and [5-$^{13}$C]glutamine to palmitate in A549 cells. Error bars, 95% confidence intervals from model (**c, d**); $*P < 0.05$. **e,** Huh7 cell proliferation after 4 days in the presence or absence of glutamine. Error bars, s.e.m. ($n = 3$) (**a, b, e**). $**P < 0.005$ comparing glutamine-free cultures. $***P < 0.001$ comparing normoxia and hypoxia.

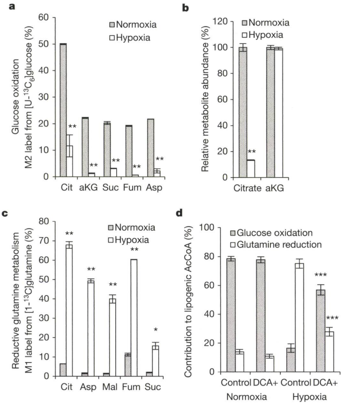


**Figure 3 | Reductive TCA metabolism increases under hypoxia.** MRC5 cells were cultured under normoxia or hypoxia for 3 days in the presence of tracer. **a,** Relative level of glucose oxidation as determined by M2 labelling from [U-$^{13}$C$_6$]glucose (see Supplementary Fig. 16a for atom transition map). M2 isotopologues were the most abundant labelled metabolites in mass spectra. **b,** Relative abundance of citrate and αKG. **c,** Relative contribution of reductive glutamine metabolism to TCA metabolites, determined by M1 labelling from [1-$^{13}$C]glutamine. **d,** Contribution of glucose oxidation and glutamine reduction to lipogenesis in A549 cells cultured with or without 5 mM DCA. Error bars, s.e.m. (**a–c**) ($n = 3$) and 95% confidence intervals (**d**). $*P < 0.01$ and $**P < 0.001$ comparing normoxia to hypoxia. $***P < 0.05$ comparing control with DCA in hypoxia.

and abundances of TCA cycle metabolites in cells grown under hypoxia. Using [U-$^{13}$C$_6$]glucose, we observed a significant decrease in relative flux through the pyruvate dehydrogenase (PDH) complex (Fig. 3a and Supplementary Fig. 16a, b). In addition, the citrate pool became depleted, which would be expected to increase reductive carboxylation flux through mass action (Fig. 3b and Supplementary Fig. 16c). On the other hand, we detected increased amounts of isotopic label in TCA cycle intermediates when using labelled glutamine tracers under hypoxia (Fig. 3c and Supplementary Fig. 16d, e). Reductively metabolized glutamine accounted for as much as 40–70% of the intracellular citrate, aspartate, malate and fumarate pools when cells were cultured in low oxygen (Fig. 3c). Given the marked reduction in PDH flux observed in hypoxia, we tested the ability of dichloroacetate (DCA) to restore PDH activity and mitigate the contribution of reductive metabolism to lipogenesis. DCA inhibits pyruvate dehydrogenase kinases (PDKs)[21], and PDK1 is a known target of HIF-1α that inhibits the activity of PDH through phosphorylation[22,23]. Although DCA treatment had no observable effect on carbon use under normoxia, reductive glutamine metabolism was inhibited and glucose oxidation was partly restored in A549 cells cultured with DCA under hypoxia (Fig. 3d and Supplementary Fig. 16f), suggesting that hypoxia-induced PDK1 expression contributes to the use of reductive carboxylation for fatty-acid synthesis.

The von Hippel–Lindau (VHL) tumour suppressor protein is frequently lost in renal cell carcinoma (RCC) and results in a state of 'pseudohypoxia' by activating HIF signalling[24–26]. To understand further the role of this pathway in promoting the switch to reductive TCA metabolism we tested RCC cells deficient in VHL using ISA. Remarkably, VHL-deficient RCC cell lines preferentially used reductive glutamine metabolism for lipogenesis, even when cultured under normal oxygen levels, whereas those expressing wild-type (WT) VHL behaved similarly to other carcinoma cell lines (Fig. 4a and Supplementary Fig. 17). Re-expression of WT VHL in previously VHL-deficient cell lines resulted in a switch back to oxidative glucose metabolism as the source of carbon for lipid synthesis (Fig. 4b), reduced extracellular fluxes of glucose, lactate and glutamine (Supplementary Fig. 18a), and increased the pool of intracellular citrate relative to αKG (Supplementary Fig. 18b) under normoxia. Furthermore, shRNA-mediated knockdown of HIF-2α partly restored glucose-mediated lipogenesis in 786-O cells (Fig. 4c, d). Consistent with VHL and HIF influencing the switch to reductive glutamine metabolism during hypoxia, glucose entry into the TCA cycle by PDH was increased under normoxia upon introduction of WT VHL or knockdown of HIF-2α in 786-O cells (Supplementary Fig. 18c). Similar changes were observed after knockdown of the HIF-α dimerization partner ARNT (aryl hydrocarbon receptor nuclear translocator) in VHL-deficient normoxic UMRC2 cells, which express both HIF-1α and HIF-2α (ref. 27), or after ARNT knockdown in hypoxic A549 and 143B cells (Supplementary Fig. 19).

Our results highlight an important role for reductive TCA metabolism of glutamine in cell proliferation at physiological oxygen levels

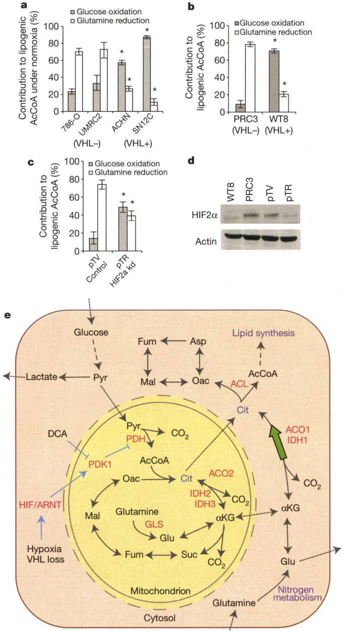


**Figure 4 | HIF/ARNT/VHL signalling regulate carbon use for lipogenesis.** **a–c**, Contribution of glucose oxidation ([U-$^{13}C_6$]glucose) and glutamine reduction ([5-$^{13}$C]glutamine) to lipogenesis in RCC lines (**a**), parental control (PRC3) and VHL+ (WT8) cells derived from 786-O line (**b**) or vector control (pTV) or HIF2α shRNA (pTR) cells derived from 786-O line (**c**). **d**, Western blot to determine HIF-2α levels for cells in **b**, **c**. Error bars, 95% confidence intervals obtained from ISA model. *$P < 0.05$. **e**, Model depicting the metabolic reprograming of mammalian cells by hypoxia or VHL loss to use reductive glutamine metabolism for lipogenesis. HIF stabilization drives transcription of PDK1, which decreases PDH activity and subsequently intracellular citrate levels. IDH1 and ACO1 reductively generate lipogenic citrate from glutamine-derived αKG. DCA can inhibit PDKs, forcing increased glucose oxidation in hypoxic cells.

(Fig. 4e). Given the almost exclusive use of reductive carboxylation for lipogenesis under hypoxia, a redundant or contributing role of mitochondrial IDH2 in this pathway is probable; however, our data provide evidence that the reductive pathway involves IDH1-mediated catalysis in the cytoplasm. Although the carbon source that cells use for lipid synthesis appears to be determined, at least partly, by HIF-mediated regulation of PDK1, additional hypoxia-associated changes may also promote reductive glutamine metabolism. For example, HIF-2α enhances c-MYC activity[28], which in turn drives glutamine catabolism through the regulation of numerous genes including glutaminase[29]. This metabolic reprogramming provides an effective, glucose-independent means of generating AcCoA for biosynthesis. Because glucose is also delivered to cells through the vasculature, it may be limited in microenvironments with decreased oxygen availability[30]. Reductively metabolizing amino acids for lipid synthesis under these conditions would allow cells to conserve glucose for production of ribose and other biosynthetic precursors (for example, one carbon pool, hexosamines) that are not typically generated from other nutrients. Thus, reductive metabolism may allow cells to distribute available nutrients more efficiently in poorly vascularized microenvironments. These results add a new dimension to our understanding of cell metabolism and suggest potential therapeutic targets along the reductive carboxylation and glutamine catabolic pathways that could mitigate hypoxic tumour growth.

## METHODS SUMMARY

For determination of steady-state labelling of polar metabolites, cells were cultured for approximately 24 h in the presence of $^{13}$C-labelled glutamine or glucose before extraction. For experiments involving stable isotopic labelling of lipid biomass, cells were grown for approximately 3–5 days in the presence of tracer before extraction. Details of the extraction and derivatization methods are described in the Supplementary Methods. Computational determination of metabolic fluxes, confidence intervals, *de novo* lipogenesis and the contribution of tracers to fatty-acid carbon was accomplished using an in-house software package, Metran[18]. Details of the metabolic networks and gas chromatography/mass spectrometry (GC/MS) measurements used for modelling and complete results are described in Supplementary Information. The generation of cells stably expressing control shRNAs or those targeted IDH1 or IDH2 is described in Supplementary Methods; all experiments were conducted within four passages of initial selection. Hypoxic microenvironments were generated by feeding incubators with a premixed gas composed of 1% $O_2$, 5% $CO_2$ and 94% $N_2$, and $O_2$ levels were confirmed to range between 1 and 3% using a Fyrite combustion analyser. For details of recombinant IDH1 production and enzyme assays, T-cell activation, medium analysis, [5-$^{14}$C]glutamine experiment and western blotting, see Supplementary Methods.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 17 March; accepted 28 September 2011.**

**Published online 20 November 2011.**

1. Hatzivassiliou, G. *et al.* ATP citrate lyase inhibition can suppress tumor cell growth. *Cancer Cell* **8,** 311–321 (2005).
2. Wellen, K. E. *et al.* ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324,** 1076–1080 (2009).
3. Migita, T. *et al.* ATP citrate lyase: activation and therapeutic implications in non-small cell lung cancer. *Cancer Res.* **68,** 8547–8554 (2008).
4. Semenza, G. L. HIF-1: upstream and downstream of cancer metabolism. *Curr. Opin. Genet. Dev.* **20,** 51–56 (2010).
5. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl Acad. Sci. USA* **104,** 19345–19350 (2007).
6. Weinberg, F. *et al.* Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. *Proc. Natl Acad. Sci. USA* **107,** 8788–8793 (2010).
7. Des Rosiers, C. *et al.* Isotopomer analysis of citric acid cycle and gluconeogenesis in rat liver. Reversibility of isocitrate dehydrogenase and involvement of ATP-citrate lyase in gluconeogenesis. *J. Biol. Chem.* **270,** 10027–10036 (1995).
8. Yoo, H., Antoniewicz, M. R., Stephanopoulos, G. & Kelleher, J. K. Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. *J. Biol. Chem.* **283,** 20621–20627 (2008).
9. Metallo, C. M., Walther, J. L. & Stephanopoulos, G. Evaluation of $^{13}$C isotopic tracers for metabolic flux analysis in mammalian cells. *J. Biotechnol.* **144,** 167–174 (2009).
10. Lemons, J. M. *et al.* Quiescent fibroblasts exhibit high metabolic activity. *PLoS Biol.* **8,** e1000514 (2010).
11. Ward, P. S. *et al.* The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17,** 225–234 (2010).
12. Boros, L. G. *et al.* Defective RNA ribose synthesis in fibroblasts from patients with thiamine-responsive megaloblastic anemia (TRMA). *Blood* **102,** 3556–3561 (2003).

13. Munger, J. *et al.* Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nature Biotechnol.* **26,** 1179–1186 (2008).
14. Maier, K., Hofmann, U., Reuss, M. & Mauch, K. Identification of metabolic fluxes in hepatic cells from transient $^{13}$C-labeling experiments: part II. Flux estimation. *Biotechnol. Bioeng.* **100,** 355–370 (2008).
15. Kharroubi, A. T., Masterson, T. M., Aldaghlas, T. A., Kennedy, K. A. & Kelleher, J. K. Isotopomer spectral analysis of triglyceride fatty acid synthesis in 3T3–L1 cells. *Am. J. Physiol.* **263,** E667–E675 (1992).
16. Siebert, G., Carsiotis, M. & Plaut, G. W. The enzymatic properties of isocitric dehydrogenase. *J. Biol. Chem.* **226,** 977–991 (1957).
17. Sauer, U. Metabolic networks in motion: $^{13}$C-based flux analysis. *Mol. Syst. Biol.* **2,** 62 (2006).
18. Young, J. D., Walther, J. L., Antoniewicz, M. R., Yoo, H. & Stephanopoulos, G. An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnol. Bioeng.* **99,** 686–699 (2008).
19. Hartong, D. T. *et al.* Insights from retinitis pigmentosa into the roles of isocitrate dehydrogenases in the Krebs cycle. *Nature Genet.* **40,** 1230–1234 (2008).
20. Cheng, T. *et al.* Pyruvate carboxylase is required for glutamine-independent growth of tumor cells. *Proc. Natl Acad. Sci. USA* **108,** 8674–8679 (2011).
21. Bonnet, S. *et al.* A mitochondria-K$^+$ channel axis is suppressed in cancer and its normalization promotes apoptosis and inhibits cancer growth. *Cancer Cell* **11,** 37–51 (2007).
22. Kim, J. W., Tchernyshyov, I., Semenza, G. L. & Dang, C. V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metab.* **3,** 177–185 (2006).
23. Papandreou, I., Cairns, R. A., Fontana, L., Lim, A. L. & Denko, N. C. HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. *Cell Metab.* **3,** 187–197 (2006).
24. Kaelin, W. G. Jr. Molecular basis of the VHL hereditary cancer syndrome. *Nature Rev. Cancer* **2,** 673–682 (2002).
25. Ivan, M. *et al.* HIFα targeted for VHL-mediated destruction by proline hydroxylation: implications for O$_2$ sensing. *Science* **292,** 464–468 (2001).
26. Jaakkola, P. *et al.* Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O$_2$-regulated prolyl hydroxylation. *Science* **292,** 468–472 (2001).
27. Zimmer, M. *et al.* Small-molecule inhibitors of HIF-2a translation link its 5′UTR iron-responsive element to oxygen sensing. *Mol. Cell* **32,** 838–848 (2008).
28. Gordan, J. D., Bertout, J. A., Hu, C. J., Diehl, J. A. & Simon, M. C. HIF-2α promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. *Cancer Cell* **11,** 335–347 (2007).
29. Gao, P. *et al.* c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458,** 762–765 (2009).
30. Sutherland, R. M. Cell and environment interactions in tumor microregions: the multicell spheroid model. *Science* **240,** 177–184 (1988).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank N. Vokes and P. Ward for discussions. We also thank S. Gross and Agios Pharmaceuticals for providing the IDH1 construct. We acknowledge support from National Institutes of Health grant R01 DK075850-01. C.M.M. is supported by a postdoctoral fellowship from the American Cancer Society. K.H. is supported by the German Research Foundation (DFG) grant HI1400. L.G. is supported by the NIH and the Glenn Foundation for Medical Research. M.G.V.H. is supported by the Burrough's Wellcome Fund, the Smith Family, the Damon Runyon Cancer Research Foundation and the National Cancer Institute. D.J.I. is an investigator of the Howard Hughes Medical Institute. O.I. is supported by R01 CA122591 and the Dana Farber/Harvard Cancer Center Kidney SPORE Grant Developmental Award.

**Author Contributions** C.M.M., P.A.G., E.L.B., K.R.M., J.Y., K.H. and C.M.J. performed cellular experiments and isotope tracing. C.M.M. and P.A.G. performed metabolite profiling and analysed data. K.R.M. and M.G.V.H. performed enzyme assays and $^{14}$C experiments. E.L.B., J.Y. and Z.R.J. generated western blots. D.J.I. and L.G. provided support and reagents. J.K.K., M.G.V.H., O.I. and G.S. provided conceptual advice. C.M.M., J.K.K., M.G.V.H., O.I. and G.S. wrote and edited the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.S. (gregstep@mit.edu) or O.I. (iliopoul@helix.mgh.harvard.edu).

doi:10.1038/nature10642

# Reductive carboxylation supports growth in tumour cells with defective mitochondria

Andrew R. Mullen[1], William W. Wheaton[2,3], Eunsook S. Jin[4,5], Pei-Hsuan Chen[1], Lucas B. Sullivan[2,3], Tzuling Cheng[1], Youfeng Yang[6], W. Marston Linehan[6], Navdeep S. Chandel[2,3] & Ralph J. DeBerardinis[1,7,8]

**Mitochondrial metabolism provides precursors to build macromolecules in growing cancer cells[1,2]. In normally functioning tumour cell mitochondria, oxidative metabolism of glucose- and glutamine-derived carbon produces citrate and acetyl-coenzyme A for lipid synthesis, which is required for tumorigenesis[3]. Yet some tumours harbour mutations in the citric acid cycle (CAC) or electron transport chain (ETC) that disable normal oxidative mitochondrial function[4–7], and it is unknown how cells from such tumours generate precursors for macromolecular synthesis. Here we show that tumour cells with defective mitochondria use glutamine-dependent reductive carboxylation rather than oxidative metabolism as the major pathway of citrate formation. This pathway uses mitochondrial and cytosolic isoforms of $NADP^+$/NADPH-dependent isocitrate dehydrogenase, and subsequent metabolism of glutamine-derived citrate provides both the acetyl-coenzyme A for lipid synthesis and the four-carbon intermediates needed to produce the remaining CAC metabolites and related macromolecular precursors. This reductive, glutamine-dependent pathway is the dominant mode of metabolism in rapidly growing malignant cells containing mutations in complex I or complex III of the ETC, in patient-derived renal carcinoma cells with mutations in fumarate hydratase, and in cells with normal mitochondria subjected to acute pharmacological ETC inhibition. Our findings reveal the novel induction of a versatile glutamine-dependent pathway that reverses many of the reactions of the canonical CAC, supports tumour cell growth, and explains how cells generate pools of CAC intermediates in the face of impaired mitochondrial metabolism.**

We first studied the metabolism of isogenic 143B human osteosarcoma cells that contained or lacked a loss-of-function mutation in ETC complex III (cytochrome *b*–c1 complex). These cell lines were generated by depleting 143B mitochondrial DNA (mtDNA) and repopulating with either wild-type mtDNA or mtDNA containing a frameshift mutation in the gene encoding cytochrome *b* (*CYTB*: also known as *MT-CYB*), an essential complex III component[8]. Despite lack of respiration and complex III function in the mutants[8], both wild-type (WT 143B) and *CYTB*-mutant (*CYTB* 143B) cells form colonies in soft agar[9] and proliferate at comparable rates (Supplementary Fig. 1a), making these cells a good model in which to study growth during mitochondrial dysfunction. Both cell lines had detectable CAC intermediates, although citrate was less abundant and succinate was significantly more abundant in the *CYTB* 143B cells (Supplementary Fig. 1b). As expected for cells with defective oxidative phosphorylation, *CYTB* 143B cells had higher glucose consumption and lactate production than WT 143B cells, indicating a metabolic shift towards aerobic glycolysis (Fig. 1a). To determine the effects of *CYTB* mutation on the metabolic fates of glucose, we cultured both cell lines in medium containing D[U-$^{13}$C]glucose (here U indicates uniformly labelled) and measured $^{13}$C enrichment of intracellular metabolites by mass spectrometry. In WT 143B cells, most citrate molecules contained glucose-derived $^{13}$C (Fig. 1b). Citrate m+2 (citrate containing two additional mass units from $^{13}$C) results from oxidative decarboxylation of glucose-derived pyruvate by pyruvate dehydrogenase (PDH) to form [1,2-$^{13}$C]acetyl-coA (coA, coenzyme A), followed by condensation with an unlabelled

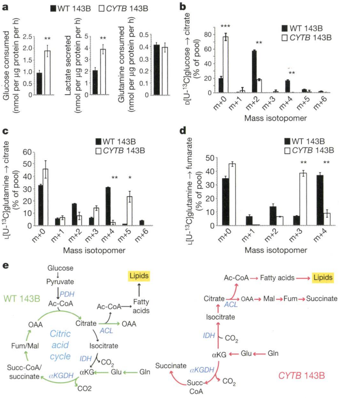


**Figure 1 | A reductive pathway of glutamine metabolism in cancer cells lacking activity of ETC complex III.** **a**, Glucose utilization, lactate secretion and glutamine utilization in WT 143B and *CYTB* 143B cells. **b**, Mass isotopomer analysis of citrate in cells cultured with D[U-$^{13}$C]glucose and unlabelled glutamine. **c**, **d**, Mass isotopomer analysis of citrate and fumarate in cells cultured with L[U-$^{13}$C]glutamine and unlabelled glucose. Data are the average ± s.d. for three independent cultures. $^*P < 0.05$; $^{**}P < 0.005$, Student's *t*-test. **e**, Schematic of glutamine metabolism in WT 143B and *CYTB* 143B cells. Coloured arrows follow the paths of glutamine-derived carbon. Abbreviations: Ac-CoA, acetyl-CoA; OAA, oxaloacetate; Gln, glutamine; Glu, glutamate; αKG, α-ketoglutarate; Succ-CoA, succinyl-CoA; Fum, fumarate; Mal, malate; PDH, pyruvate dehydrogenase; ACL, ATP-citrate lyase; IDH, isocitrate dehydrogenase; αKGDH, α-ketoglutarate dehydrogenase.

---

[1]Department of Pediatrics, University of Texas – Southwestern Medical Center at Dallas, Dallas, Texas 75390-9063, USA. [2]Department of Medicine, Northwestern University, Chicago, Illinois 60611-3008, USA. [3]Department of Cell and Molecular Biology, Northwestern University, Chicago, Illinois 60611-3008, USA. [4]Department of Internal Medicine, University of Texas – Southwestern Medical Center at Dallas, Dallas, Texas 75390-9063, USA. [5]Advanced Imaging Research Center, University of Texas – Southwestern Medical Center at Dallas, Dallas, Texas 75390-8568, USA. [6]Urologic Oncology Branch, National Cancer Institute, Bethesda, Maryland 20892, USA. [7]McDermott Center for Human Growth and Development, University of Texas – Southwestern Medical Center at Dallas, Dallas, Texas 75390-8591, USA. [8]Harold C. Simmons Comprehensive Cancer Center, University of Texas – Southwestern Medical Center at Dallas, Dallas, Texas 75235-5303, USA.

oxaloacetate (OAA). Processing of citrate m+2 around one turn of the CAC produces citrate m+4 (Supplementary Fig. 2). In *CYTB* 143B cells, most citrate contained no glucose carbon (m+0), indicating suppressed PDH contribution to acetyl-CoA (Fig. 1b). Fumarate and malate m+2 were also decreased (Supplementary Fig. 1c, d).

Glutamine is a major respiratory substrate in cancer cells, providing energy and carbon for growth[3,10]. Such cells use glutamine to supply anaplerosis, the replenishment of CAC intermediates withdrawn from the cycle to feed biosynthetic pathways. Both WT and *CYTB* 143B cells require glutamine for colony formation, implying a respiration-independent function for glutamine in cell growth[9], and both cell types consumed glutamine at similar rates (Fig. 1a). We cultured both cell lines with L[U-$^{13}$C]glutamine to define the metabolic fates of glutamine. Like other glutamine-dependent cancer cells[11], WT 143B cells used glutamine as the major anaplerotic precursor, resulting in a large amount of fumarate, malate and citrate m+4 (Fig. 1c, d; Supplementary Figs 1e and 3, top). In contrast, *CYTB* 143B cells produced only trace quantities of citrate m+4. Instead, they produced citrate m+5 through reductive carboxylation of glutamine-derived α-ketoglutarate (Fig. 1c). This reaction involves addition of an unlabelled carbon by isocitrate dehydrogenase (IDH) acting in reverse relative to the canonical oxidative CAC[12]. Cleavage of citrate m+5 by ATP-citrate lyase then produces acetyl-CoA m+2 and OAA m+3, with the unlabelled carbon retained on OAA (Supplementary Fig. 3, bottom). Examination of other CAC metabolites in *CYTB* 143B cells revealed that they were formed downstream of citrate m+5 and OAA m+3. These cells contained abundant malate and fumarate m+3, which was essentially absent from WT 143B cells (Fig. 1d and Supplementary Fig. 1e). Thus these metabolites are generated from reductive, rather than oxidative, glutamine metabolism (Fig. 1e). Succinate was formed through both reductive and oxidative glutamine metabolism in *CYTB* 143B cells (Supplementary Fig. 1f). Glutamine-dependent reductive carboxylation was also observed in *CYTB* 143B cells cultured in medium containing a full complement of amino acids (Supplementary Fig. 4), and in cells with a genetic defect in ETC complex I (Supplementary Fig. 5).

To determine which of the three IDH isoforms contributed to reductive carboxylation, we used RNA interference (RNAi) to silence expression of *IDH1* (isocitrate dehydrogenase 1 (NADP$^+$), soluble), *IDH2* and *IDH3* in *CYTB* 143B cells and measured glutamine-derived $^{13}$C labelling in citrate. Silencing either of the NADP$^+$/NADPH-dependent IDH isoforms, *IDH1* or *IDH2*, reduced citrate m+5 in cells cultured with L[U-$^{13}$C]glutamine and silencing both at once further reduced citrate m+5 (Fig. 2a, b). Conversion of glutamine to glutamate was unaffected (Fig. 2b). Silencing *IDH3*, the mitochondrial NAD$^+$/NADH-dependent isoform, had no effect on labelling (Fig. 2a, b). Thus reductive carboxylation in *CYTB* 143B cells involves *IDH1* and *IDH2*, but not *IDH3*. Because IDH1 and IDH2 proteins are localized primarily in the cytoplasm and mitochondria, respectively, of both cell lines (Supplementary Fig. 6), the data suggest that reductive carboxylation in *CYTB* 143B cells may occur in either compartment.

To determine whether IDH1 and IDH2 are required for growth, we stably silenced each enzyme using lentiviral RNAi. In WT 143B cells, silencing either isoform reduced cell growth (Supplementary Fig. 7), consistent with findings in other cancer cells with normal mitochondria[13]. In *CYTB* 143B cells, a similar growth reduction occurred when either isoform was silenced (Supplementary Fig. 7). Thus both IDH1 and IDH2 support growth in cells that use glutamine-dependent reductive carboxylation.

We next determined whether reductive glutamine metabolism contributed to lipogenesis, a biosynthetic activity required for cancer cell growth. Both WT and *CYTB* 143B cells produced fatty acids *de novo*, as indicated by incorporation of $^3H_2O$ into lipids (Fig. 3a). The two cell lines were then cultured with $^{14}$C-labelled lipogenic substrates to examine carbon preferences for lipogenesis (Fig. 3a). Both cell lines had an intact pathway of lipogenesis from exogenously supplied acetate. The enhanced labelling in *CYTB* 143B cells probably resulted from an increased specific activity of the intracellular two-carbon pool because of suppressed PDH. Lipids derived from *CYTB* 143B cells contained little radioactivity from glucose, but abundant radioactivity from glutamine, indicating a switch in the extent to which these nutrients supplied lipogenesis. To quantify the contribution of glucose and glutamine to lipogenic acetyl-CoA, cells were cultured in medium

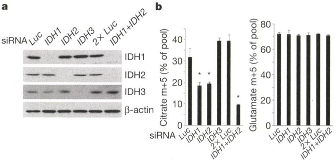


**Figure 2 | NADP$^+$/NADPH-dependent isoforms of isocitrate dehydrogenase contribute to reductive carboxylation.** **a**, Transient silencing of isocitrate dehydrogenase (IDH) proteins in *CYTB* 143B cells with short interfering RNAs (siRNAs) directed against *IDH1*, *IDH2* or *IDH3*. An siRNA directed against luciferase (*Luc*) was used as a negative control. When *IDH1* and *IDH2* siRNAs were transfected concurrently (*IDH1*+*IDH2*), a transfection with the same nanomolar amount of Luc siRNA (2×*Luc*) was used as a negative control. **b**, Mass isotopomer analysis of citrate and glutamate in *CYTB* 143B cells cultured with L[U-$^{13}$C]glutamine after silencing of IDH isoforms. Data are the average ± s.d. of three independent cultures. *$P < 0.05$, Student's *t*-test.

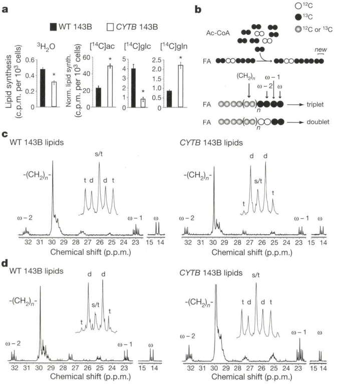


**Figure 3 | Glutamine is the major lipogenic precursor in cells lacking oxidative CAC function.** **a**, Cells were cultured in medium containing $^3H_2O$ (21 vol.%), or $^{14}$C tracers of acetate (ac, 5 μCi), glucose (glc, 10 μCi) or glutamine (gln, 10 μCi). Lipids were analysed for $^3$H or $^{14}$C content by scintillation counting. In the $^{14}$C experiments, raw counts per $10^3$ cells were normalized to the absolute rate of fatty acid synthesis established using $^3H_2O$. *$P < 0.05$, Student's *t*-test. **b**, Schematic outlining synthesis of fatty acids (FA) from acetyl-CoA (Ac-CoA) and the source of the triplet and doublet at ω − 1 in $^{13}$C NMR spectroscopy. **c**, **d**, $^{13}$C NMR of lipids labelled with D[U-$^{13}$C]glucose (**c**) or L[U-$^{13}$C]glutamine (**d**). Insets are expansions of the ω − 1 resonance. Abbreviations: s, singlet; d, doublet; t, triplet.

containing either D[U-$^{13}$C]glucose and unlabelled glutamine, or L[U-$^{13}$C]glutamine and unlabelled glucose. Extracted lipids were analysed by $^{13}$C NMR. Because fatty acids are synthesized by the sequential addition of two-carbon units, analysing $^{13}$C–$^{13}$C coupling between the $\omega - 1$ and $\omega - 2$ carbons—which enter the fatty acyl chain on different two-carbon units—enables calculation of $^{13}$C enrichment in lipogenic acetyl-CoA[3]. In the $\omega - 1$ multiplet, $^{13}$C–$^{13}$C coupling produces a triplet when two $^{13}$C-labelled two-carbon units are added in succession, whereas one labelled and one unlabelled unit produce a doublet (Fig. 3b). Therefore $^{13}$C labelling in lipogenic acetyl-CoA is proportional to the contribution of the triplet to the overall area of the $\omega - 1$ multiplet. In WT 143B cells cultured with D[U-$^{13}$C]glucose, the $\omega - 1$ resonance was dominated by the triplet, which accounted for 65% of the total area (Fig. 3c, left). Glutamine was a minor source of fatty acyl carbon, as the triplet accounted for only 15% of the area when the cells were cultured with L[U-$^{13}$C]glutamine (Fig. 3d, left). This pattern was reversed in the *CYTB* 143B cells, with glutamine providing 67% of the lipogenic carbon (Fig. 3c, d, right; Supplementary Fig. 8a). Thus, reductive glutamine metabolism was the major pathway of fatty acid production in newly synthesized lipids during *CYTB* 143B growth.

Although most cancer cells have functional mitochondria, a subset of human tumours harbours mutations that impair mitochondrial metabolism. Two major classes of mutations occur in genes required for function of the CAC enzymes succinate dehydrogenase (*SDHA, SDHB, SDHC, SDHD, SDHAF2*) or fumarate hydratase (*FH*)[14]. To test whether glutamine-dependent reductive carboxylation occurs in tumour cells harbouring these mutations, we studied UOK262 cells derived from a renal tumour in a patient with hereditary leiomyomatosis renal cell carcinoma syndrome. In this syndrome, affected individuals inherit one loss-of-function *FH* allele and their tumours display loss of heterozygosity for the *FH* locus[15]. UOK262 cells are defective in respiration and devoid of FH activity[15]. They failed to survive culture in glucose-free medium containing galactose and glutamine (Fig. 4a), a condition that forces cells to use mitochondrial metabolism to generate ATP[16]. Lack of FH activity precludes the use of glutamine as a carbon source in the conventional oxidative CAC. Yet the UOK262 cells required glutamine to proliferate (Fig. 4b). Culture of these cells with D[U-$^{13}$C]glucose produced mass isotopomer data similar to *CYTB* 143B cells, with minimal formation of citrate m+2 (Fig. 4c, left). In cells cultured with L[U-$^{13}$C]glutamine, the most abundant mass isotopomer of citrate was m+5 formed by reductive carboxylation (Fig. 4c, left). Fumarate was predominantly formed through oxidative glutamine metabolism, producing fumarate m+4 (Fig. 4c, middle). However, a large fraction of the malate pool (m+3) was formed via citrate cleavage and subsequent reductive metabolism (Fig. 4c, right). Lipid labelling using D[U-$^{13}$C]glucose or L[U-$^{13}$C]glutamine revealed that, like *CYTB* 143B cells, glutamine is the major carbon source for fatty acid synthesis (Supplementary Fig. 8b). Thus these cancer cells produce CAC intermediates using two distinct pathways to compensate for the loss of FH enzyme activity: oxidative metabolism of glutamine to fumarate, and glutamine-dependent reductive carboxylation to produce citrate, acetyl-CoA for lipogenesis and remaining four-carbon CAC intermediates.

Finally, we tested whether glutamine-dependent reductive carboxylation was unique to cells with permanent abnormalities in mitochondrial metabolism. WT 143B cells and transformed mouse embryonic fibroblasts (MEFs) were subjected to acute ETC inhibition. Exposure to inhibitors of complex III (antimycin) or complex I (rotenone) induced a switch from oxidative to reductive metabolism of glutamine (Fig. 4d, Supplementary Fig. 9a). Enhanced reductive carboxylation was evident one hour after ETC inhibition and was not eliminated by pre-treatment with the protein synthesis inhibitor cycloheximide (Supplementary Fig. 9b, c). Metformin, an antidiabetic agent that inhibits respiration through effects on complex I[17], also induced glutamine-dependent reductive carboxylation at high doses (Fig. 4d, Supplementary Fig. 9a). Interestingly, doses that mimicked plasma concentrations in patients with acute metformin toxicity and lactic acidosis induced reductive glutamine metabolism (Supplementary Fig. 10). These data demonstrate that glutamine-dependent reductive carboxylation is a common cellular response to impaired mitochondrial metabolism.

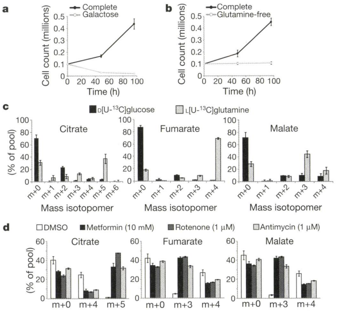


**Figure 4 | Glutamine-dependent reductive carboxylation in renal carcinoma cells lacking fumarate hydratase and during ETC inhibition in tumour cells with normal mitochondria.** Growth of UOK262 renal carcinoma cells in complete medium, in medium in which glucose was replaced with galactose (**a**), and in medium lacking glutamine (**b**). **c**, Mass isotopomer analysis of citrate, fumarate and malate from UOK262 cells cultured in medium containing D[U-$^{13}$C]glucose and unlabelled glutamine, or L[U-$^{13}$C]glutamine and unlabelled glucose. **d**, Mass isotopomer analysis for citrate, fumarate and malate from WT 143B cells cultured in medium containing unlabelled glucose and L[U-$^{13}$C]glutamine. The labelled glutamine was introduced at the same time the cells were treated with vehicle (DMSO), metformin, rotenone or antimycin. Metabolites were extracted 6 h later. In all experiments, data are the average ± s.d. of three independent cultures.

Glutamine-dependent reductive carboxylation has been described previously as a minor source of isocitrate/citrate and lipogenic carbon in mammalian cells[12,13,18–20]. In *Plasmodium falciparum*, a unicellular protozoan with limited respiration, the pathway is a major source of citrate and acetyl units[21]. We show that this pathway can also function as the predominant metabolic strategy to produce lipids for cancer cell growth, and can be induced acutely during ETC inhibition. We speculate that reductive carboxylation is stimulated by a disturbance in the redox ratio of the mitochondria caused by ETC impairment, decreasing the NAD$^+$/NADH ratio and rendering oxidative function of the CAC less efficient. This ratio could be dissipated in part by NAD(P)-transhydrogenases, which transfer reducing equivalents from NADH to NADPH, and this in turn may drive NADPH-dependent reductive carboxylation by IDH1 and IDH2. This redirection of the CAC extends the versatility of glutamine metabolism in cancer cell growth. In cells with normal mitochondria, glutamine oxidation is frequently the major source of anaplerosis, providing most of the carbon to pools of CAC intermediates and facilitating production of macromolecular precursors in the mitochondria[22]. Glutamine-dependent reductive carboxylation enables glutamine to retain this status as a growth-promoting nutrient even when mitochondrial metabolism is impaired. Therefore, during malignant transformation the acquisition of somatic mutations that impair mitochondrial function and stimulate aerobic glycolysis (the Warburg effect) does not diminish the importance of glutamine metabolism or of the mitochondria themselves for tumour cell proliferation.

## METHODS SUMMARY

WT 143B, *CYTB* 143B, UOK262, CCL16-B2 and CCL16-NDI1 cells were cultured as described[9,15,23]. MEFs were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS, Hyclone), penicillin/streptomycin and 6 mM L-glutamine. Cell growth was monitored in sub-confluent cultures by trypsinizing and counting with a haemacytometer. Glucose, lactate and glutamine were measured as described[24]. Isotopic labelling was performed in DMEM with 10% dialysed FBS supplemented with either 15 mM D[U-$^{13}$C]glucose or 2 mM L[U-$^{13}$C]glutamine. Aqueous metabolites were analysed using an Agilent 6970 gas chromatograph and an Agilent 5973 mass selective detector. To determine relative metabolite abundance across samples, the area of the total ion current peak for the relevant metabolite was compared to that of an internal standard (2-oxobutyrate) and normalized for protein content. Analysis of $^{13}$C enrichment and mass isotopomer distribution was performed as published[11,25]. For radioisotope lipid synthesis assays, cells were cultured for 18 h in complete medium supplemented with $^3H_2O$ or $^{14}$C tracers (Perkin Elmer). Lipids were extracted and analysed as described[3]. For lipid synthesis experiments using NMR, cells were plated in 10-cm dishes and allowed to proliferate exponentially in medium containing D[U-$^{13}$C]glucose and unlabelled glutamine, or unlabelled glucose and L[U-$^{13}$C]glutamine until two confluent 15-cm dishes were obtained. Lipids were then extracted and analysed by $^{13}$C NMR as described[3]. Gene silencing was performed using commercial small interfering RNAs (Thermo) directed against *IDH1*, *IDH2* or *IDH3*. Cells were transfected with DharmaFECT transfection reagent (Thermo) and protein abundance was examined after 72 h using western blots with commercial antibodies (Santa Cruz Biotechnology). Cells were incubated with L[U-$^{13}$C]glutamine for two hours before extraction of metabolites for mass isotopomer distribution analysis. Stable silencing used lentiviral shRNAs from the Mission shRNA pLKO.1-puro library (Sigma).

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 13 June; accepted 13 October 2011.**

**Published online 20 November 2011.**

1. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–1033 (2009).
2. Gao, P. *et al.* c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458,** 762–765 (2009).
3. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl Acad. Sci. USA* **104,** 19345–19350 (2007).
4. Linehan, W. M., Srinivasan, R. & Schmidt, L. S. The genetic basis of kidney cancer: a metabolic disease. *Nature Rev. Urol.* **7,** 277–285 (2010).
5. Baysal, B. E. *et al.* Mutations in SDHD, a mitochondrial complex II gene, in hereditary paraganglioma. *Science* **287,** 848–851 (2000).
6. Tomlinson, I. P. *et al.* Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nature Genet.* **30,** 406–410 (2002).
7. Hao, H. X. *et al.* SDH5, a gene required for flavination of succinate dehydrogenase, is mutated in paraganglioma. *Science* **325,** 1139–1142 (2009).
8. Rana, M., de Coo, I., Diaz, F., Smeets, H. & Moraes, C. T. An out-of-frame cytochrome b gene deletion from a patient with parkinsonism is associated with impaired complex III assembly and an increase in free radical production. *Ann. Neurol.* **48,** 774–781 (2000).
9. Weinberg, F. *et al.* Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. *Proc. Natl Acad. Sci. USA* **107,** 8788–8793 (2010).
10. Reitzer, L. J., Wice, B. M. & Kennell, D. Evidence that glutamine, not sugar, is the major energy source for cultured HeLa cells. *J. Biol. Chem.* **254,** 2669–2676 (1979).
11. Cheng, T. *et al.* Pyruvate carboxylase is required for glutamine-independent growth of tumor cells. *Proc. Natl Acad. Sci. USA* **108,** 8674–8679 (2011).
12. Des Rosiers, C., Fernandez, C. A., David, F. & Brunengraber, H. Reversibility of the mitochondrial isocitrate dehydrogenase reaction in the perfused rat liver. Evidence from isotopomer analysis of citric acid cycle intermediates. *J. Biol. Chem.* **269,** 27179–27182 (1994).
13. Ward, P. S. *et al.* The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17,** 225–234 (2010).
14. Frezza, C., Pollard, P. J. & Gottlieb, E. Inborn and acquired metabolic defects in cancer. *J. Mol. Med.* **89,** 213–220 (2011).
15. Yang, Y. *et al.* UOK 262 cell line, fumarate hydratase deficient (FH-/FH-) hereditary leiomyomatosis renal cell carcinoma: *in vitro* and *in vivo* model of an aberrant energy metabolic pathway in human cancer. *Cancer Genet. Cytogenet.* **196,** 45–55 (2010).
16. Rossignol, R. *et al.* Energy substrate modulates mitochondrial structure and oxidative capacity in cancer cells. *Cancer Res.* **64,** 985–993 (2004).
17. El-Mir, M. Y. *et al.* Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. *J. Biol. Chem.* **275,** 223–228 (2000).
18. Yoo, H., Antoniewicz, M. R., Stephanopoulos, G. & Kelleher, J. K. Quantifying reductive carboxylation flux of glutamine to lipid in a brown adipocyte cell line. *J. Biol. Chem.* **283,** 20621–20627 (2008).
19. Comte, B., Vincent, G., Bouchard, B., Benderdour, M. & Des Rosiers, C. Reverse flux through cardiac NADP(+)-isocitrate dehydrogenase under normoxia and ischemia. *Am. J. Physiol. Heart Circ. Physiol.* **283,** H1505–H1514 (2002).
20. Lemons, J. M. *et al.* Quiescent fibroblasts exhibit high metabolic activity. *PLoS Biol.* **8,** e1000514 (2010).
21. Olszewski, K. L. *et al.* Branched tricarboxylic acid metabolism in *Plasmodium falciparum*. *Nature* **466,** 774–778 (2010).
22. DeBerardinis, R. J. & Cheng, T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* **29,** 313–324 (2009).
23. Brunelle, J. K. *et al.* Oxygen sensing requires mitochondrial ROS but not oxidative phosphorylation. *Cell Metab.* **1,** 409–414 (2005).
24. Yang, C. *et al.* Glioblastoma cells require glutamate dehydrogenase to survive impairments of glucose metabolism or Akt signaling. *Cancer Res.* **69,** 7986–7993 (2009).
25. Fernandez, C. A., Des Rosiers, C., Previs, S. F., David, F. & Brunengraber, H. Correction of $^{13}$C mass isotopomer distributions for natural stable isotope abundance. *J. Mass Spectrom.* **31,** 255–262 (1996).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** K. Uyeda and members of the DeBerardinis and Chandel laboratories provided critical assessment of the data, and J. Sudderth and C. Yang provided experimental assistance. We thank C. Moraes and I.F.M. de Coo for the WT 143B and *CYTB* 143B cell lines, I.E. Scheffler for CCL16-B2 cells and T. Yagi for CCL16-NDI1 cells. This work was supported by grants to R.J.D. from the NIH (K08DK072565 and R01CA157996), the Cancer Prevention and Research Institute of Texas (CPRIT, HIRP100437) and the Robert A. Welch Foundation (I1733); to N.S.C. from the NIH (R01CA123067), the LUNGevity Foundation and a Consortium of Independent Lung Health Organizations convened by Respiratory Health Association of Metropolitan Chicago; and to E.S.J. by an NIH grant (DK078933). The work was also supported by the Intramural Research Program of the NIH, National Cancer Institute Center for Cancer Research and by NIH grant RR02584. NIH training grants supported A.R.M. (5T32GM083831), W.W.W. (T32CA009560) and L.B.S. (T32GM008061). T.C. was supported by a CPRIT training grant.

**Author Contributions** A.R.M., W.W.W., N.S.C. and R.J.D. designed the research. A.R.M., W.W.W., L.B.S., E.S.J., T.C. and P.-H.C. performed the experiments. A.R.M., W.W.W., L.B.S., E.S.J., P.-H.C., T.C., N.S.C. and R.J.D. analysed the data. Y.Y. and W.M.L. provided the FH-deficient (UOK262) cells. A.R.M., N.S.C. and R.J.D. wrote the paper.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.J.D. (ralph.deberardinis@utsouthwestern.edu) or N.S.C. (nav@northwestern.edu).

doi:10.1038/nature10730

# Differential oestrogen receptor binding is associated with clinical outcome in breast cancer

Caryn S. Ross-Innes[1], Rory Stark[1], Andrew E. Teschendorff[2], Kelly A. Holmes[1], H. Raza Ali[1,8], Mark J. Dunning[1], Gordon D. Brown[1], Ondrej Gojis[3,4,5], Ian O. Ellis[6], Andrew R. Green[6], Simak Ali[3], Suet-Feung Chin[1], Carlo Palmieri[3], Carlos Caldas[1,7,8,9] & Jason S. Carroll[1,7]

**Oestrogen receptor-α (ER) is the defining and driving transcription factor in the majority of breast cancers and its target genes dictate cell growth and endocrine response, yet genomic understanding of ER function has been restricted to model systems[1–3]. Here we map genome-wide ER-binding events, by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq), in primary breast cancers from patients with different clinical outcomes and in distant ER-positive metastases. We find that drug-resistant cancers still recruit ER to the chromatin, but that ER binding is a dynamic process, with the acquisition of unique ER-binding regions in tumours from patients that are likely to relapse. The acquired ER regulatory regions associated with poor clinical outcome observed in primary tumours reveal gene signatures that predict clinical outcome in ER-positive disease exclusively. We find that the differential ER-binding programme observed in tumours from patients with poor outcome is not due to the selection of a rare subpopulation of cells, but is due to the FOXA1-mediated reprogramming of ER binding on a rapid timescale. The parallel redistribution of ER and FOXA1 binding events in drug-resistant cellular contexts is supported by histological co-expression of ER and FOXA1 in metastatic samples. By establishing transcription-factor mapping in primary tumour material, we show that there is plasticity in ER-binding capacity, with distinct combinations of *cis*-regulatory elements linked with the different clinical outcomes.**

Recent technological advances have allowed mapping of ER-binding events, with the goal of discovering the *cis*-regulatory elements and factors involved in mediating ER binding and transcription. Several genome-wide maps of ER in breast cancer cell-line models exist[1–3], all showing that most ER-binding events occur at distal *cis*-regulatory elements. Forkhead motifs are enriched within the regions bound by ER binding and multiple studies have identified the forkhead protein FOXA1 as an important pioneer factor for ER–chromatin interactions[4–6]. However, the ER mapping studies have been restricted to breast cancer cell lines, mostly the MCF-7 cell line. We sought to interrogate ER-binding events, for the first time, in primary frozen breast cancer samples, to determine if ER binding is dynamic and if specific *cis*-regulatory elements can distinguish tumours from patients with distinct clinical outcomes.

ER ChIP-seq was performed in eight ER$^+$, progesterone receptor (PR)$^+$, HER2$^-$ primary breast tumours, representative of tumours from patients with better prognosis[7], a conclusion supported by the available long-term clinical follow-up (Supplementary Fig. 1). Also included were seven primary breast tumours from patients with a poor outcome (ER$^+$ PR$^-$ HER2$^-$ or ER$^+$ PR$^+$ HER$^+$), because PR$^-$ or HER2$^+$ tumours are more likely to be aggressive[8,9]. As expected, the poor outcome patients who had long-term clinical follow-up died of breast cancer (Supplementary Fig. 1). Furthermore, three ER$^+$ distant metastatic samples from women with breast cancer were included. The metastatic locations and sample preparation can be found in Supplementary Fig. 1. As a control, we included two breast cancer samples that were ER$^-$ (ER-α negative), but expressed high transcript levels of ER-β.

ER ChIP-seq was conducted and ER-binding peaks were called using two different algorithms, MACS[10] and SWEMBL (http://www.ebi.ac.uk/~swilder/SWEMBL/), to minimize peak caller bias. The number of sequencing reads and ER-binding events for each tumour is shown in Supplementary Fig. 2. ER binding could be mapped in all tumours, but total peak intensity and the number of identified binding events differed. Three tumours were split into two sections and ER ChIP-seq was conducted on the separate sections. We found very good concordance when comparing different sections of the same tumour ($R^2 = 0.954$) suggesting that tumour heterogeneity did not substantially influence the ER-binding signal obtained from a sample (Supplementary Fig. 3).

We initially assessed whether a conserved set of breast cancer ER-binding events could be identified. We found a core set of 484 ER-binding events that were identified in at least 75% of all the tumours, but not in either of the ER$^-$ tumours (Fig. 1a). Peak calling details can be found in Supplementary Fig. 4. An example of a core ER-binding event is shown in Fig. 1b. This demonstrates that ER binding to chromatin still occurs even in tumours that are unlikely to respond to anti-oestrogen therapies, implying that drug resistance is not due to loss of ER binding to DNA. The average ER-binding signal intensity was highest in the metastatic samples and lowest in patients with good outcome tumours, a phenomenon observed both within the 484 core ER-binding regions (Fig. 1c) and globally (Supplementary Fig. 4). These data indicate that there is an acquisition of binding signal intensity in tumours that progress towards a poorer prognosis and ultimately metastasize. The only DNA motif found to be enriched in the core ER-binding events was an oestrogen responsive element (ERE) (Fig. 1d). The genes near (within 20 kilobases (kb): an optimal window between ER-binding events and target genes[11]) the 484 core ER-binding events exhibited elevated expression in the ER$^+$ tumours used for ChIP-seq, as compared to all other genes (data not shown) and were higher in ER$^+$ tumours relative to ER$^-$ tumours in nine independent data sets (Supplementary Fig. 5). The genes are provided in Supplementary Fig. 6 and include classic ER target genes such as *TFF1*, *GREB1* and *RARA*. A gene predictor was generated based on genes near the core ER-binding events. Patients were stratified and the tumours with the highest 'risk index' had a poor clinical outcome when compared to the tumours with the lowest 'risk index' (Fig. 1e shows the results based on one study[12] and additional data sets are shown in Supplementary Fig. 6; only ER$^+$ patients were considered). These conserved *cis*-regulatory elements

---

[1] Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. [2] UCL Cancer Institute, University College London, 72 Huntley Street, London WC1E 6BT, UK. [3] Imperial College London, Hammersmith Campus, London W12 0NN, UK. [4] Department of Gynaecology and Obstetrics, Third Faculty of Medicine, Charles University, Ruska 87, Prague 10, 100 00, Czech Republic. [5] Department of Pathology, Third Faculty of Medicine, Charles University, Ruska 87, Prague 10, 100 00, Czech Republic. [6] Department of Histopathology, Nottingham University Hospitals NHS Trust, Nottingham City Hospital, Nottingham NG5 1PB, UK. [7] Department of Oncology, University of Cambridge CB2 0XZ, UK. [8] Cambridge Breast Unit, Addenbrooke's Hospital, Cambridge University Hospital NHS Foundation Trust and NIHR Cambridge Biomedical Research Centre, Cambridge CB2 2QQ, UK. [9] Cambridge Experimental Cancer Medicine Centre (ECMC), Cambridge CB2 0RE, UK.

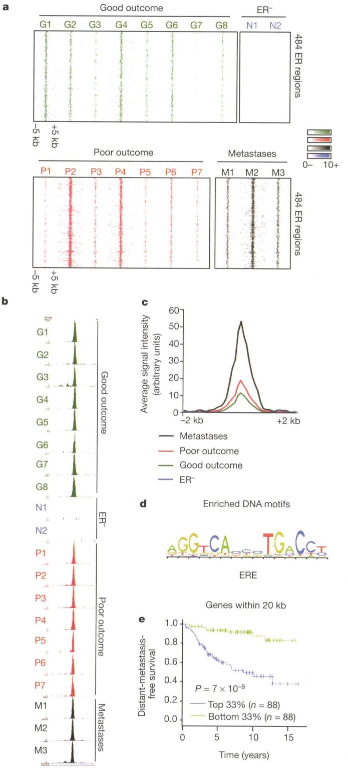


**Figure 1 | A subset of ER-binding events is conserved in primary breast tumours and distant metastases.** **a**, Heat map showing binding peak intensity of 484 core ER-binding events that are common to primary breast tumours and distant metastases. The window represents ±5 kb regions from the centre of the binding events. **b**, Example of an ER-binding event at the *RARA* locus, which is present in all primary ER$^+$ breast tumours and metastases, but not the ER$^-$ tumours. **c**, Normalized average signal intensity of all core ER-binding events. Also included are the ER$^-$ tumours. **d**, Motif analysis revealed the enrichment of EREs. **e**, Genes within 20 kb of the core ER-binding events were used to generate a gene predictor that was tested in independent data sets for predictive value. Tumours were stratified according to expression of this gene signature and the top one third and bottom one third of tumours were compared. Results from one data set are shown[12] and additional data sets are provided in Supplementary Fig. 6.

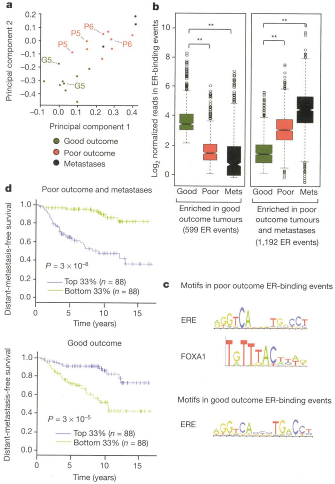


**Figure 2 | ER-binding profiles can discriminate between tumours from patients with different clinical outcomes.** **a**, Principal component analysis (44% of total variance) of the 1,791 ER-binding events that can discriminate between the patients with good outcome tumours and those with poor/met tumours. Included are the replicates from three tumours, which are highlighted. **b**, Box plot representing distribution of normalized read counts in differential ER-binding events that are statistically enriched in either the patients with good outcome tumours (599 ER-binding events) or the patients with poor outcome tumours and metastases (Mets) (1,192 ER-binding events). The samples were pre-normalized. $^{**}P < 1 \times 10^{-10}$. *P* values were calculated using the Wilcoxon rank sum test. **c**, Enriched motifs in the poor/met and the good outcome ER-binding events. **d**, Genes within 20 kb of the differentially bound ER-binding events were used to generate a gene predictor. Tumours were stratified according to expression of this gene signature and the top one third and bottom one third of tumours were compared. Results from one data set are shown[12] and additional data sets are provided in Supplementary Fig. 12.

and their putative target genes may be the elements that contribute to tumorigenesis and are maintained regardless of the clinical outcome of the breast cancer patient. In contrast to the primary breast cancers and metastases, we mapped ER binding in three normal human mammary glands and two normal human liver samples and found limited numbers of ER-binding events, with almost no concordance in ER binding between individuals (data not shown).

We sought to determine if differential ER-binding events could discriminate the patients with good outcome (ER$^+$ PR$^+$ HER2$^-$ tumours), from patients with poor outcome or metastases (we described the combined set as poor/met tumours). After normalization of the data to account for global differences in ER binding, differential binding analysis (DBA) was used to identify ER-binding events that were statistically enriched in one category or the other. This resulted in a set of ER-binding events that could discriminate between the two groups when using principal component analysis

(Fig. 2a). In total, DBA revealed 1,192 genomic regions that had significantly more ER binding in the poor/met group compared to the good outcome patients (Fig. 2b), and 599 ER binding regions with more ER binding in the good outcome patients when compared to the poor/met patients (Fig. 2b). The clustering of the tumours on the basis of the 1,791 differential ER-binding events can be visualized in Supplementary Fig. 7. These findings suggest that there are specific and re-occurring *cis*-regulatory elements that are occupied by ER in breast cancers, but that these are different in tumours that respond to treatment versus those that relapse and metastasize. Analysis of enriched DNA motifs identified the presence of ERE and FOXA1 motifs in the differential poor outcome ER-binding events and ERE motifs in the good outcome ER-binding events (Fig. 2c). Correlation of the poor outcome ER-binding events with known processes revealed an association with endocrine resistance and luminal B status (Supplementary Fig. 8).

To investigate if the genes near the differential ER-binding events were potentially functional in breast cancer, we analysed genes within a 20 kb window around the 1,192 poor/met and 599 good outcome ER-binding events. Using a training set, we generated a gene expression predictor for each of the good and poor outcome gene lists. The probability calculation and comparisons between the good and poor outcome genes is shown in Supplementary Figs 9 and 10. Within the poor outcome gene list was the oncogene *ERBB2* (all genes are shown in Supplementary Fig. 11). As expected, genes in the poor outcome predictor were preferentially upregulated in poor outcome patients, whereas those in the good outcome predictor were preferentially downregulated (Supplementary Fig. 10). We next tested the predictors in an independent large cohort of breast cancer patients[12], only considering ER$^+$ tumours. Using distant-metastasis-free survival as an endpoint, both gene sets predicted outcome ($P = 3 \times 10^{-5}$ for good and $3 \times 10^{-8}$ for poor outcome genes) in this data set[12] and with the expected opposite directionality (Fig. 2d). The gene predictors were associated with survival in additional data sets (Supplementary Fig. 12) and were largely independent of histopathological factors (Supplementary Fig. 13). We tested 1,000 randomizations from the entire list of genes and determined that the probability that a random set of genes would yield an equally robust predictor of clinical outcome was $P = 0.004$. Furthermore, the good and poor gene predictors had no predictive power in four cohorts of ER$^-$ patients (Supplementary Fig. 14). This suggests that the increased ER binding at distinct *cis*-regulatory elements is functionally and biologically relevant, resulting in altered gene expression profiles that contribute to differences in drug response and overall survival.

To validate the findings made in the tumours, we explored the possibility that ER-binding events were acquired in cell-line models of endocrine resistance. ER binding was mapped by ChIP-seq in three commonly used tamoxifen-responsive, ER$^+$ breast cancer cell lines (MCF-7, T-47D and ZR75-1), and two tamoxifen-resistant, ER$^+$ breast cancer cell lines, namely a tamoxifen-resistant MCF-7 derivative (TAM-R)[13] and BT-474 cells that are ER$^+$ and contain the *ERBB2* amplification (ER$^+$ HER2$^+$). Similar to BT474 cells, TAM-R cells have increased ERBB2 protein levels[13] and both represent cellular systems in which increased growth factor signalling results in endocrine resistance. For all five cell lines, ER ChIP-seq was performed in at least duplicate, in asynchronous cells to recapitulate the situation observed in primary tumours (Supplementary Fig. 2).

Almost 7,000 (6,920) ER-binding events were identified in all replicates of all cell lines (Fig. 3a and Supplementary Fig. 15). Almost all (98.9%) of the core ER-binding events that occurred in most primary tumours (Fig. 1a) overlapped with the cell-line core ER-binding events. DBA identified 8,188 ER-binding events with significantly stronger binding affinity in the tamoxifen-resistant cell lines, and 5,713 ER-binding events that were stronger in the tamoxifen-responsive cell lines (Fig. 3a). Examples of differentially bound regions are shown in Fig. 3b. Using the differential ER-binding events, the cell-line classification can be visualized by principal component analysis (Fig. 3c) and hierarchical clustering (Fig. 3d). Enriched motif analysis revealed ERE and FOXA1 motifs in regions showing increased ER binding in tamoxifen-resistant cell lines (Fig. 3e), which are the same motifs observed in the poor outcome ER-binding events in primary tumours (Fig. 2c). GATA motifs were enriched in ER-binding events depleted during acquisition of drug resistance (Fig. 3e), possibly due to competition between FOXA1 and GATA3, another prominent breast cancer transcription factor.

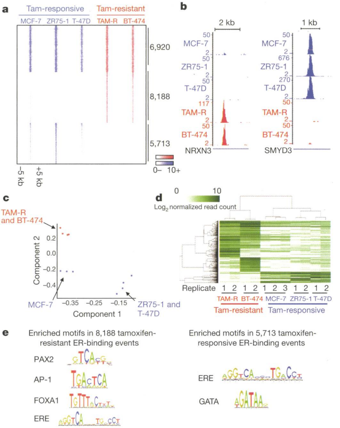


**Figure 3 | Identification of a tamoxifen-resistant ER-binding profile.** **a**, Heat map representing ER-binding events found in all tamoxifen (Tam)-responsive and tamoxifen-resistant cell lines, or those enriched in either responsive or resistant cell lines. The window represents ±5 kb regions from the centre of the binding events. Numbers indicate the total number of ER-binding events in each category. **b**, Examples of ER-binding events that are unique to either tamoxifen-responsive or tamoxifen-resistant cell lines. **c**, Principal component analysis (65% of total variance) of differential ER-binding events observed in responsive (blue) or resistant (red) breast cancer cells. **d**, Hierarchical clustering of the top 1,500 differentially bound ER-binding events (FDR $<0.006$). **e**, Enriched motifs within the ER-binding events that discriminate between tamoxifen-sensitive and resistant cell lines.

We hypothesized that the ER-binding events induced in the tamoxifen-resistant breast cancer cell lines would be the same regions that were enriched in the poor/met clinical samples. However, 79.8% of the 1,192 ER binding events enriched in the poor/met samples (from Fig. 2b) overlap with ER-binding events in wild-type MCF-7 cells, suggesting that the cell-line models are closer to the tumours and metastases from poor outcome patients. In support of this, the 599 good outcome ER-binding events observed in primary tumours (Fig. 2b) overlap poorly with the ER binding observed in MCF-7 cells (30.2% versus 79.8% for the poor outcome). Interestingly, MCF-7 cells (plus T-47D and ZR75-1 cells) were derived from the pleural effusion of metastatic breast cancer patients, but were established before tamoxifen use in the clinic. We propose that MCF-7, ZR75-1 and T-47D cell lines possess an intermediate ER-binding profile with the acquisition of additional ER-binding regions required for resistance to anti-oestrogen treatment.

The differences in ER binding between responsive and resistant contexts may be due to selection and expansion of a resistant subpopulation, or may be due to reprogramming of ER binding following specific stimuli. Growth factor pathways have long been implicated in modulating endocrine response[14,15] and have been shown to influence ER-binding potential and gene expression profiles[14]. We identified various stimuli (EGF, IL-6, TNF-α and IGF-I) shown to induce increased cellular invasion and drug resistance and treated asynchronous MCF-7 cells with control or the cocktail of mitogens for 90 min (Fig. 4a). Duplicate ER ChIP-seq replicates were performed (Supplementary Fig. 2).

Differential binding analysis identified 6,089 ER-binding regions that were differentially enriched (≥4 fold change difference, false discovery rate (FDR) <0.1) following a 90 min treatment with the cocktail. These mitogen-induced ER-binding differences could be visualized using principal component analysis (Fig. 4b). Because FOXA1 is a pioneer factor required for ER–chromatin interactions[4] and FOXA1 motifs were enriched in both the mitogen-induced ER-binding events (Supplementary Fig. 16) and the tumours from the poor outcome patients (Fig. 2c), we assessed whether the rapid, reprogrammed ER binding occurred at regions pre-determined by FOXA1. We repeated the mitogen treatment, but mapped FOXA1 binding by ChIP-seq and found that ~25% of the reprogrammed ER-binding events (1,515) occur at regions that are already bound by FOXA1, before mitogen treatment (Fig. 4c). A substantial proportion (37.6%) of the other reprogrammed ER-binding events occur at regions where FOXA1 binding is also induced by mitogens. As such, ~53% of mitogen-induced ER-binding events occur at regions pre-bound by FOXA1 or at regions that also acquire FOXA1 binding (Fig. 4c), a level of concordance that mirrors the ~50% overlap observed in wild-type cells[6].

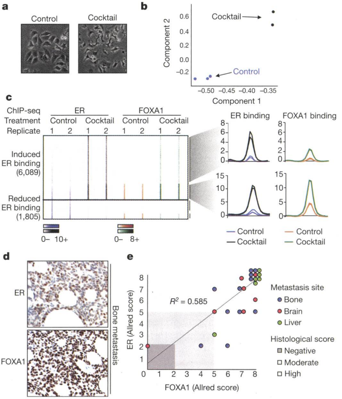


**Figure 4 | ER and FOXA1 binding is dynamic and their expression correlates in metastases. a**, Morphological changes in MCF-7 cells treated with a cocktail of mitogens (EGF, IGF-1, IL-6 and TNF-α) for 90 min. Magnification, ×40. **b**, Principal component analysis (93% of total variance) of differentially bound ER-binding events can discriminate between the control and mitogenic cocktail-treated samples in the first (and second) component. **c**, Heat map showing induction and loss of ER and FOXA1 binding after 90 min of treatment with the mitogenic cocktail. Also shown is the normalized average signal intensity of all ER- and FOXA1-binding events within the 'gained' ER-binding events. **d**, Metastatic samples from different sites were stained by immunohistochemistry for ER and FOXA1. Example of ER and FOXA1 expression in a breast cancer bone metastasis. Magnification, ×40. **e**, Graphical representation of ER and FOXA1 co-expression in metastases. Location of metastases is provided, as is the Allred score.

To determine if FOXA1 expression was present in $ER^+$ distant metastases, we obtained 24 metastatic samples (bone, brain and liver) from $ER^+$ breast cancer patients and performed immunohistochemistry for ER and FOXA1 (Fig. 4d). We found that ~87% of the metastases retained ER expression and that FOXA1 expression occurred in ~95% of the metastases (Fig. 4e). Importantly, the concordance between ER and FOXA1 was high ($R^2 = 0.585$), regardless of the site of metastasis. Therefore, the co-expression of ER and FOXA1 in distant metastases supports our conclusions that FOXA1 mediates ER reprogramming.

By mapping ER binding in clinical samples, we provide a first glimpse of the primary regulatory regions that contribute to differences within $ER^+$ breast cancers, rather than secondary events such as gene expression profiles. Our findings suggest that there is plasticity in ER binding, with distinct ER-binding profiles associated with clinical outcome. These differential ER-binding profiles seem to be mediated by FOXA1. A remaining question is what dictates differential FOXA1 and subsequently ER binding. Possibilities include changes in the genomic landscape, alterations in co-factor levels or changes in FOXA1 structure and function, potentially by post-translational modifications. By establishing transcription factor mapping in primary samples, we find that differential ER-binding patterns govern gene expression programs and are associated with clinical outcome in $ER^+$ cancer.

## METHODS SUMMARY

MCF-7, ZR75-1, T-47D and BT-474 human cell lines were obtained from the American type culture collection (ATCC) and grown in the relevant media. TAM-R cells[13] were a gift from I. Hutcheson and R. Nicholson (Cardiff). The $ER^+$ breast cancer tumours were obtained from the Nottingham Tenovus primary breast cancer series, Addenbrooke's Hospital and Imperial College Healthcare NHS Trust, with appropriate ethical approval from the repositories. The malignant pericardial effusion and the two distant metastases were obtained from Imperial College Healthcare NHS Trust. For ChIP in the tumours and metastases, the frozen sample was cut into smaller pieces before ChIP, which was then performed as previously described[16]. For the malignant pericardial effusion, epithelial cells were first enriched using Dynabeads conjugated with Epcam[17]. For ChIPs from cell-line material, proliferating cells were cross-linked and processed for ChIP as previously described[16]. The antibodies used were anti-ER (sc-543) from Santa Cruz Biotechnologies and anti-FOXA1 (ab5089) from Abcam. Sequences generated by the Illumina Genome Analyzer were processed by the Illumina analysis pipeline version 1.6.1, and aligned to the Human Reference Genome (assembly hg18, NCBI build 36.1, March 2008) using BWA version 0.5.5 (ref. 18). Differential binding analysis was performed using the DiffBind package[19]. For immunohistochemical analyses, ER staining was conducted using the 6F11/2 mouse monoclonal antibody (Novocastra, Leica Microsystems) and FOXA1 staining was conducted using a rabbit polyclonal antibody (ab23738) from Abcam. An Allred scoring system was used to assess staining accounting for both staining intensity and the proportion of cells stained.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 19 May; accepted 23 November 2011.**

**Published online 4 January 2012.**

1. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nature Genet.* **38,** 1289–1297 (2006).
2. Lin, C. Y. *et al.* Whole-genome cartography of estrogen receptor α binding sites. *PLoS Genet.* **3,** e87 (2007).
3. Welboren, W. J. *et al.* ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* **28,** 1418–1428 (2009).

4. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122,** 33–43 (2005).
5. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132,** 958–970 (2008).
6. Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature Genet.* **43,** 27–33 (2011).
7. EBCTCG. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* **365,** 1687–1717 (2005).
8. Kun, Y. *et al.* Classifying the estrogen receptor status of breast cancers by expression profiles reveals a poor prognosis subpopulation exhibiting high expression of the ERBB2 receptor. *Hum. Mol. Genet.* **12,** 3245–3258 (2003).
9. Arpino, G. *et al.* Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J. Natl Cancer Inst.* **97,** 1254–1261 (2005).
10. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).
11. Fullwood, M. J. *et al.* An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462,** 58–64 (2009).
12. Loi, S. *et al.* Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J. Clin. Oncol.* **25,** 1239–1246 (2007).
13. Knowlden, J. M. *et al.* Elevated levels of epidermal growth factor receptor/c-erbB2 heterodimers mediate an autocrine growth regulatory pathway in tamoxifen-resistant MCF-7 cells. *Endocrinology* **144,** 1032–1044 (2003).
14. Lupien, M. *et al.* Growth factor stimulation induces a distinct ERα cistrome underlying breast cancer endocrine resistance. *Genes Dev.* **24,** 2219–2227 (2010).
15. Nagashima, T. *et al.* Quantitative transcriptional control of ErbB receptor signaling undergoes graded to biphasic response for cell differentiation. *J. Biol. Chem.* **282,** 4045–4056 (2007).
16. Schmidt, D. *et al.* ChIP-seq: using high-throughput sequencing to discover protein–DNA interactions. *Methods* **48,** 240–248 (2009).
17. Gomm, J. J. *et al.* Isolation of pure populations of epithelial and myoepithelial cells from the normal human mammary gland using immunomagnetic separation with Dynabeads. *Anal. Biochem.* **226,** 91–99 (1995).
18. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
19. Stark, R. & Brown, G. D. DiffBind: differential binding analysis of ChIP-seq peak data. *Bioconductor* http://bioconductor.org/packages/release/bioc/html/DiffBind.html.

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** The authors would like to thank D. Schmidt for assistance with figures, J. Hadfield for Illumina sequencing, S. MacArthur, O. Rueda, S. Vowler, R. Russell and M. Wilson for technical and bioinformatics help. We thank J. Stingl and his laboratory for help with the normal mammary gland work. We would like to acknowledge the support of The University of Cambridge, Cancer Research UK and Hutchison Whampoa Limited. The authors would like to thank Imperial College Healthcare NHS Trust, Human Biomaterials Resource Centre (Tissue Bank). Tumour samples from Cambridge were obtained with support from NIHR Biomedical Research Centre and the Experimental Cancer Medicine Centre. C.S.R.-I. is supported by a Commonwealth Scholarship. O.G. is part funded by a grant awarded by the Ministry of Education of the Czech Republic (Project "Oncology" MSM 0021620808) and is also a recipient of a Translational Research Fellowship from the European Society of Medical Oncology. C.P. is funded by Cancer Research UK. J.S.C. is supported by an ERC starting grant and an EMBO Young investigator award.

**Author Contributions** C.S.R.-I., R.S., C.C. and J.S.C. designed all experiments. Experimental work was conducted by C.S.R.-I. with help from K.A.H. Computational analysis was conducted by R.S. and A.E.T., with help from M.J.D. and G.D.B. All clinical samples, clinical information and help with sample processing was provided by C.C., C.P., S.-F.C., S.A., A.R.G., I.O.E. and O.G. Histological analysis was conducted by H.R.A. The manuscript was written by C.S.R.-I., R.S., C.C. and J.S.C. with assistance from other authors.

**Author Information** Alignment and peak data are deposited in the Gene Expression Omnibus under accession number GSE32222. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.S.C. (jason.carroll@cancer.org.uk) or C.C. (carlos.caldas@cancer.org.uk).

doi:10.1038/nature10624

# Natural killer cells act as rheostats modulating antiviral T cells

Stephen N. Waggoner[1], Markus Cornberg[2], Liisa K. Selin[1] & Raymond M. Welsh[1]

**Antiviral T cells are thought to regulate whether hepatitis C virus (HCV) and human immunodeficiency virus (HIV) infections result in viral control, asymptomatic persistence or severe disease, although the reasons for these different outcomes remain unclear. Recent genetic evidence, however, has indicated a correlation between certain natural killer (NK)-cell receptors and progression of both HIV and HCV infection[1–3], implying that NK cells have a role in these T-cell-associated diseases. Although direct NK-cell-mediated lysis of virus-infected cells may contribute to antiviral defence during some virus infections—especially murine cytomegalovirus (MCMV) infections in mice and perhaps HIV in humans[4,5]—NK cells have also been suspected of having immunoregulatory functions. For instance, NK cells may indirectly regulate T-cell responses by lysing MCMV-infected antigen-presenting cells[6,7]. In contrast to MCMV, lymphocytic choriomeningitis virus (LCMV) infection in mice seems to be resistant to any direct antiviral effects of NK cells[5,8]. Here we examine the roles of NK cells in regulating T-cell-dependent viral persistence and immunopathology in mice infected with LCMV, an established model for HIV and HCV infections in humans. We describe a three-way interaction, whereby activated NK cells cytolytically eliminate activated CD4 T cells that affect CD8 T-cell function and exhaustion. At high virus doses, NK cells prevented fatal pathology while enabling T-cell exhaustion and viral persistence, but at medium doses NK cells paradoxically facilitated lethal T-cell-mediated pathology. Thus, NK cells can act as rheostats, regulating CD4 T-cell-mediated support for the antiviral CD8 T cells that control viral pathogenesis and persistence.**

Intravenous (i.v.) inoculation of C57BL/6 mice with a low ($5 \times 10^4$ plaque-forming units (p.f.u.), medium ($2 \times 10^5$ p.f.u.), or high ($2 \times 10^6$ p.f.u.) dose of LCMV, strain clone 13, resulted in different degrees of pathology, as indicated by weight loss (Fig. 1a) and by histological analysis of lung sections at day 15 post-infection (p.i.) (Fig. 1b). The high dose caused a precipitous drop in body weight during the first week of infection (Fig. 1a, right) but, thereafter, clonal exhaustion and deletion of LCMV-specific T cells resulted in a persistent infection[9,10] associated with minimal lung pathology (Fig. 1b, right) and 100% (77 of 77) survival (Fig. 1c, top). Selective depletion of NK cells using 25 μg of anti-NK1.1 monoclonal antibodies (Supplementary Fig. 1) 1 day before high-dose infection resulted in 58% (35 of 65) mortality between days 9 and 13 of infection (Fig. 1c, top) associated with severe pulmonary oedema (data not shown) and reduced viral titres by day 7 p.i. (Fig. 1d, right). Under these high-dose conditions, therefore, the presence of NK cells promoted persistence and prevented mortality.

In contrast to the beneficial role of NK cells during high-dose infection, NK-cell depletion prevented the severe weight loss (Fig. 1a, middle) and tissue pathology (Fig. 1b, middle) associated with the medium dose of LCMV. Twenty-three per cent (7 of 31) of control-treated mice succumbed to the medium dose during the second week of infection, and the lungs of surviving mice exhibited bronchus-associated lymphoid tissue, pulmonary oedema and interstitial mononuclear infiltration. Lung pathology was absent in NK-cell-depleted mice, which uniformly survived medium-dose challenge (Fig. 1c, bottom). Moreover, although high levels of replicating virus persisted in surviving control mice at day 15 p.i., NK-cell depletion resulted in complete viral clearance (Fig. 1d, middle). In this case the presence of NK cells was detrimental for the host, as they promoted immune pathology and death.

Irrespective of the presence of NK cells, inoculation with a low dose of virus was uniformly non-lethal in 18 of 18 (100%) control and 18 of 18 (100%) of NK-cell-depleted mice by >50 days p.i., with minimal weight loss (Fig. 1a, left) and minimal lung pathology (Fig. 1b, left). Virus was completely cleared in both groups of mice by day 15 of low-dose infection (data not shown), but NK-cell depletion resulted in more rapid elimination of LCMV in liver by day 7 p.i. (Fig. 1d, left).

The weight loss, lung pathology and mortality observed in medium-dose-infected wild-type mice (Fig. 1a, b) did not occur after infection of αβ T-cell-receptor-deficient (*Tcrb*$^{-/-}$)mice, and NK-cell depletion of *Tcrb*$^{-/-}$ mice did not alter weight loss or viral burden (Supplementary Fig. 2). Thus, NK cells regulate viral clearance and immunopathology during LCMV infection through a T-cell-dependent mechanism.

As early as day 6 after medium-dose infection, the proportion and number of interferon-γ (IFN-γ)$^+$ LCMV-specific CD8 T cells was increased two- to sixfold in mice depleted of NK cells (Fig. 2a and Supplementary Fig. 3), and antiviral T cells from these mice showed an enhanced ability to co-produce tumour necrosis factor (TNF) (Supplementary Fig. 3). The number of LCMV epitope NP$_{396-404}$ tetramer-binding CD8 T cells in the spleen on day 5 p.i. was increased 4- to 20-fold in NK-cell-depleted mice relative to non-depleted control mice after infection with all doses of virus (Fig. 2b). The number of virus-specific IFN-γ$^+$ CD4 T cells was also amplified 7- to 20-fold by NK-cell depletion compared to control mice on different days after medium-dose infection (Fig. 2c). Moreover, co-production of TNF and interleukin-2 (IL-2) by antiviral CD4 T cells was augmented by NK-cell depletion (Fig. 2d and Supplementary Fig. 3). The increased magnitude of the LCMV-specific T-cell response in the absence of NK cells during medium-dose infection correlated with rapid viral clearance (Fig. 2e). Depletion of NK cells using a carefully titrated dose of anti-asialo GM1 antibody, which eliminates NK cells but not CD8 T cells[11], also enhanced antiviral CD4 and CD8 T-cell responses during medium-dose infection (Supplementary Fig. 4).

The enhanced antiviral T-cell responses suggested that NK-cell depletion may augment proliferation of LCMV-specific T cells. Transfer of carboxyfluorescein diacetate succinimidyl ester (CFSE)-labelled Thy1.1$^+$ T cells revealed a larger population of CFSE$^{\text{low}}$ donor CD4 (Fig. 2f) and CD8 (data not shown) T cells in multiple host tissues at day 6 p.i. of high-dose infection in the absence of NK cells. There was also greater specific lysis of viral-peptide-coated target cells as detected by a conventional *in vivo* cytotoxicity assay at day 4 of infection (Supplementary Fig. 5). Moreover, LCMV-specific Ly5.1$^+$ TCR transgenic (P14) CD8 T cells (transfer $10^4$) were recovered from tissues of NK-cell-depleted recipient (Ly5.2$^+$) mice at two- to ninefold greater

---

[1]Department of Pathology and Program in Immunology and Virology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA. [2]Department of Gastroenterology and Hepatology, Medizinische Hochschule, Hannover 30625, Germany.

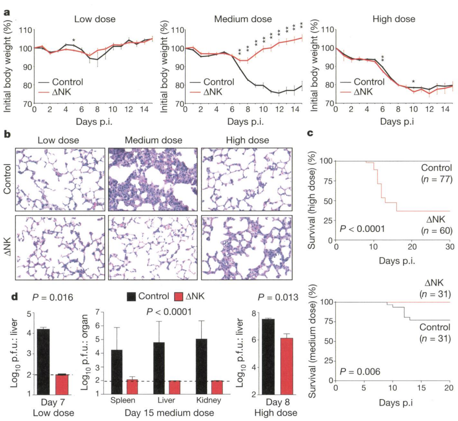


**Figure 1 | NK cells influence T-cell-dependent pathology and viral persistence during LCMV infection. a–d**, C57BL/6 mice treated with IgG2a (Control) or anti-NK1.1 (ΔNK) were infected with low ($5 \times 10^4$ p.f.u.), medium ($2 \times 10^5$ p.f.u.), or high ($2 \times 10^6$ p.f.u.) doses of LCMV. **a**, Weight loss (mean ± s.e.m.) during infection ($n = 3$–$43$ per group per day), $^*P < 0.05$, $^{**}P < 0.01$. **b**, Haematoxylin & eosin (H&E) staining of lung (×400) at day 15 p.i. **c**, Survival after high- or medium-dose infection. **d**, Viral titres (mean ± s.e.m.) after low (day 7, $n = 3$ per group), medium (day 15, $n = 9$–$15$ per group), or high (day 8, $n = 3$ per group) dose infection. Dotted line represents limit of detection.

numbers than control-treated mice 6 days after low-dose infection (Supplementary Fig. 5). Together these results indicate that NK1.1$^+$ cells repress the size of the antiviral T-cell response during LCMV infection.

The activities of CD4 T cells are important for maintaining CD8 T-cell function during LCMV infection[12–14]. To assess whether CD4 T cells were involved in the NK-cell suppression of LCMV-specific CD8 T cells, mice were treated with antibodies to concurrently deplete both NK and CD4 T cells. Whereas depletion of NK cells before medium-dose LCMV infection resulted in a >200-fold reduction in splenic viral titres at day 7 p.i. relative to control and CD4-depleted (ΔCD4) mice (Fig. 3a), depletion of both NK and CD4 T cells (ΔNKΔCD4) had no effect on viral titres. The increase in the number of antiviral CD8 T cells producing more than one cytokine caused by NK-cell depletion was also prevented by co-depletion of CD4 T cells (Fig. 3b). In contrast, co-depletion of NK and CD8 T cells did not prevent an increase in IFN-$\gamma^+$ LCMV epitope $GP_{61-80}$-specific CD4 T cells (control: $3.8 \pm 0.5\%$ versus ΔNK: $9.8 \pm 0.7\%$ versus ΔNK/ΔCD8: $10.7 \pm 1.6\%$, $n = 3$, $P < 0.05$ versus control) at day 12 of medium-dose infection.

Paradoxically, at the high virus dose, co-depletion of NK and CD4 T cells prevented the severe pulmonary oedema (Fig. 3c) and increased mortality (Fig. 3d) associated with depletion of NK cells alone. In this experiment, mice were harvested on day 12 p.i., when three surviving NK-cell-depleted mice were moribund and required euthanasia, whereas all double-depleted mice showed relatively normal vigour. The livers of NK/CD4 double-depleted mice contained 25-fold more p.f.u. than livers from mice depleted of NK cells alone (NK: $5.7 \pm 0.2$ p.f.u. versus ΔNK/ΔCD4: $7.1 \pm 0.1$ p.f.u., $n = 5$, $P < 0.0001$). Enhancement of LCMV-specific CD8 T cells in the absence of NK cells was also abrogated by concurrent depletion of CD4 T cells (Supplementary Fig. 6). Together these data indicate that CD4 T cells are needed for NK-cell modulation of antiviral CD8 T-cell responses associated with viral clearance, persistence and immunopathology.

We used a modified *in vivo* cytotoxicity assay by injecting splenocytes from medium-dose-infected NK-cell-depleted mice (Ly5.1$^+$, day 4 p.i.) into medium-dose-infected NK-cell-depleted (ΔNK) or isotype IgG2a-treated (control) recipient mice (Ly5.2$^+$, day 3 p.i.). After 5 h, similar proportions of total donor T (control: $0.16 \pm 0.03\%$ versus ΔNK: $0.15 \pm 0.02\%$, $n = 21$, $P = 0.80$) and B cells (control: $1.8 \pm 0.2\%$ versus ΔNK: $1.7 \pm 0.2\%$, $n = 21$, $P = 0.88$) were recovered from infected recipients, regardless of NK-cell depletion. Likewise, recovery of activated (CD44$^{hi}$ CD43(1B11)$^+$) donor CD8 T cells was similar from spleens of control and ΔNK mice, with minimal loss relative to uninfected control mice (Fig. 4a). In contrast, there was a substantial loss of activated donor CD4 T cells in infected relative to uninfected recipients, and this loss was prevented by depletion of NK cells (Fig. 4a). The magnitude of NK-cell-dependent loss of activated donor CD4 T cells was similar in low-, medium- and high-dose-infected recipients (Fig. 4b). More activated CD4 T cells, both donor and host derived, in infected (control) mice stained positively for the apoptosis indicator annexin V in comparison to naive donor CD4 T cells or to activated donor CD4 T cells in medium-dose-infected ΔNK recipient mice (Fig. 4c). In contrast to activated donor CD4 T cells, the recoveries of naive (CD44$^{low}$) phenotype CD4 and CD8 donor T cells were not altered by NK-cell depletion (data not shown). These data indicate that NK cells in wild-type mice selectively and rapidly target activated CD4 T cells for elimination during LCMV infection.

We next examined the involvement of NK-cell cytolytic mediators Fasl, TNF and perforin (Prf1) in this process. The loss of activated wild-type donor CD4 T cells in infected wild-type recipient mice (Fig. 4a, d) was seen when activated *Fas*$^{lpr}$ (Fas mutant) mouse donor cells were transferred into wild-type recipient mice or when wild-type donor cells were transferred into *Tnf*$^{-/-}$ recipient mice (Fig. 4d). In contrast, there was relatively little loss of activated wild-type donor CD4 T cells in *Prf1*$^{-/-}$ hosts (Fig. 4d), whose retention of activated donor CD4 cells was not significantly different ($P > 0.1$) from that in NK-cell-depleted wild-type or *Prf1*$^{-/-}$ hosts. Thus, NK-cell elimination of activated CD4 T cells is mediated through a perforin-dependent pathway that does not require Fas or TNF.

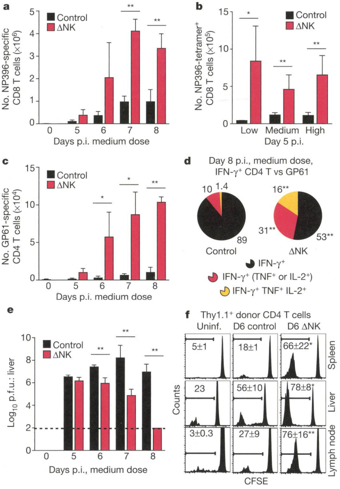


**Figure 2 | LCMV-specific T-cell responses enhanced in NK-cell-depleted mice. a–c**, Number (mean ± s.e.m.) of LCMV-specific T cells measured by IFN-γ expression (**a**, **c**) or tetramer-binding (**b**) at various days p.i. (**a**, **c**) (medium dose, $n = 3$ per group per day) or at day 5 p.i. (**b**) (various doses, $n = 3$–$11$ per group). **d**, Co-production of TNF and IL-2 by gated IFN-$\gamma^+$ CD4 T cells (day 8 p.i., medium dose) after *in vitro* stimulation with GP61 peptide. **e**, Viral titres in liver ($n = 3$–$11$ per group per day, medium dose). **f**, CFSE dilution by donor (Thy1.1$^+$) CD4 T cells in uninfected (Uninf.) and infected (medium dose, day 6 p.i.) control and ΔNK Thy1.2$^+$ host mice. Control versus ΔNK mice, $*P < 0.05$, $**P < 0.01$.

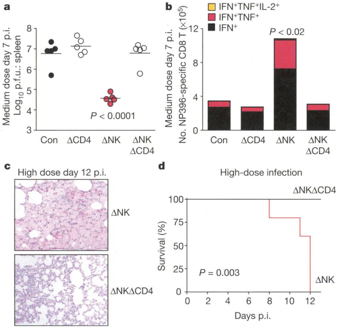


**Figure 3 | Role of CD4 T cells in NK suppression of antiviral CD8 T cells, viral control and immunopathology. a–e**, Prior to medium- (**a**, **b**) or high-dose (**c–e**) infection, mice were injected with isotype (Con, control), anti-NK1.1 (ΔNK), anti-CD4 (ΔCD4), or both (ΔNKΔCD4). **a**, **b**, Viral titres in spleen (**a**) and number of cytokine-producing NP$_{396\text{-}404}$-specific CD8 T cells ($n = 5$ per group, ± s.e.m.) in spleen at day 7 of medium-dose infection (**b**). **c**, **d**, H&E staining of lung (×100) (**c**) and survival at day 12 of high-dose infection ($n = 3$–$5$ per group) (**d**).

Previous work has implicated NKG2D (also known as Klrk1) in targeting of activated T cells by murine NK cells *in vitro*[15–17], but we observed no differences in activated wild-type donor CD4 T-cell survival in wild-type versus *Nkg2d*$^{-/-}$ recipients (Fig. 4d) or in wild-type mice treated with a blocking monoclonal antibody to NKG2D[18] (data not shown). Of note is that we did not observe the expression of ligands for activating NK-cell receptors including NKG2D, NKp46 (also known as NCR1), DNAM-1 (also known as CD226) and TRAIL (also known as TNFSF10) on these early activated CD4 T cells, and NK-cell-mediated elimination of activated donor CD4 T cells also occurred in antibody-deficient (μMT$^{-/-}$) mice (data not shown), precluding a role for antibody-dependent mechanisms. Activated CD4 T cells did, however, express much higher levels of adhesion molecules than naive cells, and these molecules have previously been shown to trigger NK-cell cytotoxicity via LFA-1[19,20]. Somewhat surprising was the observation that the activated CD4 T cells were far more susceptible than activated CD8 T cells to direct killing by the NK cells, even though both expressed high levels of adhesion molecules. We previously had shown that the presence of the negatively signalling receptor CD244 (2B4) on NK cells prevented NK-cell-mediated lysis of activated CD8 T cells[21]. We found here that although expression of the CD244 ligand CD48 was upregulated on T cells after medium-dose LCMV infection, expression levels of CD48 were much higher on activated CD8 than on activated CD4 cells (mean fluorescence intensity (MFI) of CD48: activated CD4, 3,423 ± 147; activated CD8, 6,180 ± 166; $n = 9$, $P < 0.0001$) (Supplementary Fig. 7).

To assess whether NK-cell-mediated lysis of activated CD4 T cells is a general principle of virus infections, we examined the loss of LCMV-activated CD4 T cells after transfer into mice inoculated with an unrelated *Arenavirus*, Pichinde virus (PV), the *Coronavirus* mouse hepatitis virus (MHV), or the interferon inducer and NK-cell activator polyinosinic:polycytidylic acid (polyI:C). All three stimuli induced measureable loss of activated donor CD4 T cells that was dependent upon the presence of NK cells (Fig. 4e). In reciprocal experiments, CD4 T cells activated during infection with PV, MHV, vaccinia virus (VV) or MCMV were lost upon transfer into mice infected with medium-dose LCMV when NK cells were present (Fig. 4f).

An analysis, by *in vivo* cytotoxicity assays after transfer into NK-cell-sufficient mice, of the window of time during which NK-cell regulation of T cells occurred in the LCMV medium-dose model showed reduced frequencies of activated donor CD4 T cells relative to uninfected recipients 1 (43%), 2 (32%), 3 (26%), 4 (4%) and 5 (8%) days after infection (Supplementary Fig. 8). The frequencies of activated donor CD4 T cells were increased by NK-cell depletion in recipient mice only at day 2 and day 3 p.i. These results indicate that NK cells target activated CD4 T cells mainly on the second and third day of infection, when the cytolytic activity of NK cells is at its peak[22].

These results show that NK cells can have a crucial role in controlling virus-associated morbidity, mortality and persistence in the absence of direct NK-cell-mediated control of virus replication, and they do so by altering the numbers and polyfunctionality of virus-specific T cells. Their effect on activated CD4 T cells was presumably due to direct cytotoxicity, as demonstrated by rapid perforin-dependent elimination of activated CD4 T cells by NK cells in short-term *in vivo* cytotoxicity assays and the observation of enhanced annexin reactivity

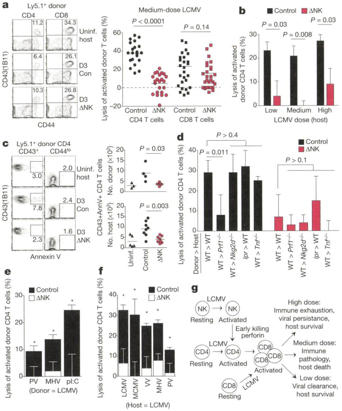


**Figure 4 | NK cells rapidly eliminate activated CD4 T cells. a–f,** *In vivo* cytotoxicity assays were performed as described in Methods using donor cells from NK-cell-depleted wild-type (Ly5.1$^+$) (**a–f**) or *Fas*$^{\text{lpr}}$ (*lpr*) (**d**) mice 4 days p.i. with medium-dose LCMV (**a–e**) or other viruses (**f**). Recovery of donor T cells was examined 5 h after transfer into control or NK-depleted (ΔNK) Ly5.2$^+$ mice (**a**, $n = 21$–28 per group; **b–f**, $n = 3$–5 per group). Host mice were uninfected or inoculated with medium-dose LCMV (day 3 p.i.) (**a**, **c**, **d**, **f**), various doses of LCMV (day 3 p.i.) (**b**), or PV (day 2), MHV (day 3) and polyI:C (pI:C) (day 1) (**e**). **a**, **b**, **d–f**, Lysis (mean ± s.e.m.) of donor T cells in infected relative to uninfected recipient mice; control versus ΔNK mice, $^*P < 0.05$. WT, wild type. **c**, Annexin V reactivity of donor (Ly5.1$^+$) and host (Ly5.1$^-$) CD4 T cells. **g**, Proposed model connecting NK-cell killing of CD4 T cells to CD8 T cells and infection outcome in the presence of NK cells.

of CD4 T cells in the presence of NK cells. The effect of activated NK cells on CD8 T cells and virus clearance seemed to be indirect, and depended on the presence of CD4 T cells, which are known to produce factors that preserve CD8 T-cell viability and functionality[12,13,23–25]. This suggests a three-way interaction, whereby NK cells suppress the CD4 T-cell response, thereby preventing augmentation of the CD8 T-cell response, which, in turn, directly regulates viral clearance and immunopathology in this system (Fig. 4f).

Our observation of direct NK-cell-mediated lysis of T cells during virus infection is distinct from published accounts of NK-cell regulation of antiviral T cells during MCMV infection, in which NK-cell-mediated lysis of virus-infected cells contributes to control of viral burden and persistence of MCMV-infected dendritic cells that in turn regulate activity of antiviral T cells[6,7,26]. Moreover, we found no NK-cell-dependent changes in the number and antigen-presenting function of splenic dendritic cells during LCMV infection (Supplementary Fig. 9), consistent with our finding that NK cells directly regulated the T cells. A possible concern was that *in vivo* cellular depletion with antibodies against NK1.1 and CD4 have the potential to affect frequencies of NKT, γδ T and regulatory T cells, but the frequencies of these lymphocytes in the spleen of LCMV-infected mice were not altered after 3 to 4 days of infection by the low concentration of anti-NK1.1 used in these studies (Supplementary Figs 1 and 10). Moreover, depletion of NK cells in γδ T-cell-deficient (*Tcrd*$^{-/-}$) and NK T-cell-deficient (*Cd1d*$^{-/-}$) mice during medium-dose infection enhanced LCMV-specific T-cell responses and reduced viral loads (Supplementary Fig. 9), similar to that in wild-type mice. Therefore, these lymphocyte lineages seem to be dispensable for NK-cell immunoregulatory function during LCMV infection.

By adjusting the dose of the LCMV inocula, it is possible to generate diverse patterns of CD8 T-cell-regulated pathogenesis, similar to the variety of pathogenic patterns a human HCV infection can take, including rapid viral clearance, severe T-cell-dependent immunopathology and long-term persistence. We show here that at a high dose of LCMV, NK cells act beneficially by suppressing T-cell responses, thereby preventing severe pathology and mortality while enabling the development of a persistent infection from which mice eventually recover and clear the virus[27]. At the medium-dose inoculum, NK-cell suppression of T cells is detrimental to the host, as virus clearance is impaired due to the limited number and functionality of T cells. However, a medium dose of virus is not sufficient for complete clonal exhaustion of T cells, ultimately resulting in severe T-cell-dependent immunopathology that can lead to death of the host.

These results indicate that NK cells can serve as rheostats, or master regulators, of antiviral T-cell responses. Consistent with the fact that many virus infections induce cytokines that potently activate NK cells[28], we found that NK-cell lysis of activated CD4 T cells was triggered by several viruses as well as after inoculation with polyI:C, which induces interferon and activates NK cells. Although a previous study found that NK-cell depletion did not alter the magnitude of antiviral T-cell responses during infection with the Armstrong strain of LCMV[29], we have observed enhanced antiviral T-cell responses and improved viral control at early time points after infection of NK-cell-depleted mice with both LCMV Armstrong and Pichinde virus (S.N.W., unpublished observations). Thus, the timing and the type of evaluation may be important to detect detrimental effects of NK cells on T cells during more benign viral infections.

## METHODS SUMMARY

**Infection model.** One day before infection, male C57BL/6 mice were selectively depleted of NK cells through a single intraperitoneal (i.p.) injection of 25 μg anti-NK1.1 monoclonal antibody (PK136) or a control mouse IgG2a (both from Bio-X-Cell), as previously described[21] (Supplementary Fig. 1). In some cases, mice were also depleted of CD4 T cells by i.p. injection of 100 μg anti-CD4 (GK1.5) at days −1 and +3 of infection. Mice were then infected i.v. with $5 \times 10^4$ (low dose), $2 \times 10^5$ (medium dose) or $2 \times 10^6$ (high dose) p.f.u. of the clone 13 variant of LCMV. Virus was titrated by plaque assay on Vero cells. In some experiments, mice were inoculated i.p. with $1.5 \times 10^7$ p.f.u. of PV, $8 \times 10^5$ p.f.u. of MHV strain A59, $1 \times 10^6$ p.f.u. of VV strain Western Reserve, $1 \times 10^6$ p.f.u. of Smith strain MCMV, or 200 μg of polyI:C (Invivogen).

**Immune assays.** The number of LCMV-specific T cells was measured by H-2D$^{\text{b}}$-NP$_{396\text{–}404}$ tetramer staining or by intracellular cytokine staining after 5 h *ex vivo* stimulation with 1 μM viral peptide in the presence of brefeldin A. T-cell cytolytic activity was measured *in vivo* as described previously[21].

***In vivo* NK-cell assay.** An unconventional *in vivo* cytotoxicity assay was previously established to determine NK-cell killing of lymphocyte populations[21]. Donor mice were depleted of NK cells and then infected i.v. or i.p. with different viruses. At day 4 p.i., single-cell splenocyte suspensions were prepared from these mice, labelled with 2 μM CFSE, and transferred ($2 \times 10^7$) into various strains of NK-cell-depleted or control recipient mice that were either uninfected or had been infected with virus 1–5 days previously. Spleens of recipient mice were harvested 5 h after transfer and assessed for survival of donor T cells.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

**Received 19 April; accepted 12 October 2011.**

**Published online 20 November 2011.**

1. Martin, M. P. *et al.* Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nature Genet.* **31,** 429–434 (2002).
2. Jennes, W. *et al.* Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. *J. Immunol.* **177,** 6588–6592 (2006).

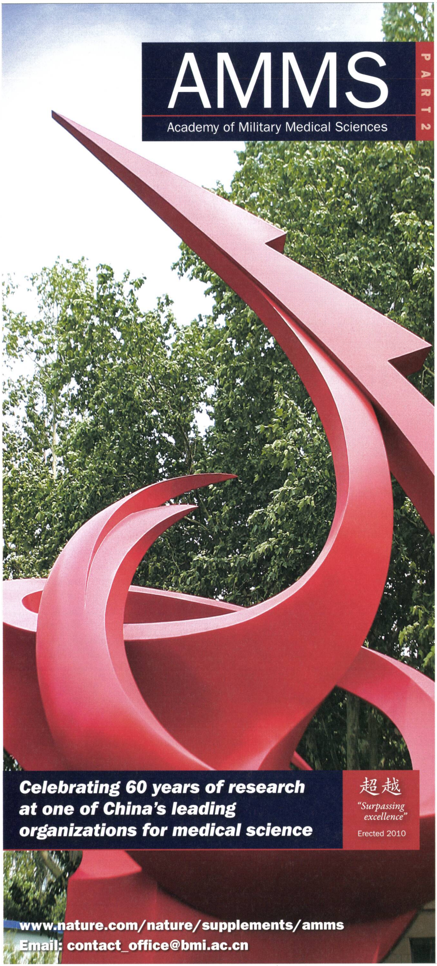
3. Khakoo, S. I. *et al.* HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science* **305,** 872–874 (2004).
4. Alter, G. *et al.* HIV-1 adaptation to NK-cell-mediated immune pressure. *Nature* **476,** 96–100 (2011).
5. Bukowski, J. F., Woda, B. A., Habu, S., Okumura, K. & Welsh, R. M. Natural killer cell depletion enhances virus synthesis and virus-induced hepatitis *in vivo*. *J. Immunol.* **131,** 1531–1538 (1983).
6. Robbins, S. H. *et al.* Natural killer cells promote early CD8 T cell responses against cytomegalovirus. *PLoS Pathog.* **3,** e123 (2007).
7. Andrews, D. M. *et al.* Innate immunity defines the capacity of antiviral T cells to limit persistent infection. *J. Exp. Med.* **207,** 1333–1343 (2010).
8. Welsh, R. M., Brubaker, J. O., Vargas-Cortes, M. & O'Donnell, C. L. Natural killer (NK) cell response to virus infections in mice with severe combined immunodeficiency. The stimulation of NK cells and the NK cell-dependent control of virus infections occur independently of T and B cell function. *J. Exp. Med.* **173,** 1053–1063 (1991).
9. Ahmed, R., Salmi, A., Butler, L. D., Chiller, J. M. & Oldstone, M. B. Selection of genetic variants of lymphocytic choriomeningitis virus in spleens of persistently infected mice. Role in suppression of cytotoxic T lymphocyte response and viral persistence. *J. Exp. Med.* **160,** 521–540 (1984).
10. Zajac, A. J. *et al.* Viral immune evasion due to persistence of activated T cells without effector function. *J. Exp. Med.* **188,** 2205–2213 (1998).
11. Yang, H., Yogeeswaran, G., Bukowski, J. F. & Welsh, R. M. Expression of asialo GM1 and other antigens and glycolipids on natural killer cells and spleen leukocytes in virus-infected mice. *Nat. Immun. Cell Growth Regul.* **4,** 21–39 (1985).
12. Battegay, M. *et al.* Enhanced establishment of a virus carrier state in adult CD4$^+$ T-cell-deficient mice. *J. Virol.* **68,** 4700–4704 (1994).
13. Matloubian, M., Concepcion, R. J. & Ahmed, R. CD4$^+$ T cells are required to sustain CD8$^+$ cytotoxic T-cell responses during chronic viral infection. *J. Virol.* **68,** 8056–8063 (1994).
14. Fuller, M. J., Khanolkar, A., Tebo, A. E. & Zajac, A. J. Maintenance, loss, and resurgence of T cell responses during acute, protracted, and chronic viral infections. *J. Immunol.* **172,** 4204–4214 (2004).
15. Noval Rivas, M. *et al.* NK cell regulation of CD4 T cell-mediated graft-versus-host disease. *J. Immunol.* **184,** 6790–6798 (2010).
16. Rabinovich, B. A. *et al.* Activated, but not resting, T cells can be recognized and killed by syngeneic NK cells. *J. Immunol.* **170,** 3572–3576 (2003).
17. Soderquest, K. *et al.* Cutting edge: CD8$^+$ T cell priming in the absence of NK cells leads to enhanced memory responses. *J. Immunol.* **186,** 3304–3308 (2011).
18. Ogasawara, K. *et al.* NKG2D blockade prevents autoimmune diabetes in NOD mice. *Immunity* **20,** 757–767 (2004).
19. Naganuma, H. *et al.* Increased susceptibility of IFN-$\gamma$-treated neuroblastoma cells to lysis by lymphokine-activated killer cells: participation of ICAM-1 induction on target cells. *Int. J. Cancer* **47,** 527–532 (1991).
20. Barber, D. F., Faure, M. & Long, E. O. LFA-1 contributes an early signal for NK cell cytotoxicity. *J. Immunol.* **173,** 3653–3659 (2004).
21. Waggoner, S. N., Taniguchi, R. T., Mathew, P. A., Kumar, V. & Welsh, R. M. Absence of mouse 2B4 promotes NK cell-mediated killing of activated CD8$^+$ T cells, leading to prolonged viral persistence and altered pathogenesis. *J. Clin. Invest.* **120,** 1925–1938 (2010).
22. Welsh, R. M. Jr. Cytotoxic cells induced during lymphocytic choriomeningitis virus infection of mice. I. Characterization of natural killer cell induction. *J. Exp. Med.* **148,** 163–181 (1978).
23. Yi, J. S., Du, M. & Zajac, A. J. A vital role for interleukin-21 in the control of a chronic viral infection. *Science* **324,** 1572–1576 (2009).
24. Frohlich, A. *et al.* IL-21R on T cells is critical for sustained functionality and control of chronic viral infection. *Science* **324,** 1576–1580 (2009).
25. Elsaesser, H., Sauer, K. & Brooks, D. G. IL-21 is required to control chronic viral infection. *Science* **324,** 1569–1572 (2009).
26. Bukowski, J. F., Woda, B. A. & Welsh, R. M. Pathogenesis of murine cytomegalovirus infection in natural killer cell-depleted mice. *J. Virol.* **52,** 119–128 (1984).
27. Oldstone, M. B. Biology and pathogenesis of lymphocytic choriomeningitis virus infection. *Curr. Top. Microbiol. Immunol.* **263,** 83–117 (2002).
28. Welsh, R. M. Regulation of virus infections by natural killer cells. *Nat. Immun. Cell Growth Regul.* **5,** 169–199 (1986).
29. Su, H. C. *et al.* NK cell functions restrain T cell responses during viral infections. *Eur. J. Immunol.* **31,** 3048–3055 (2001).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** We thank K. Hearn, C. Baer, J. Suschak and P. Afriyie for technical support; K. Daniels and M. Seedhom for insightful discussions; R. Taniguchi and V. Kumar for sharing unpublished observations; and H. Ducharme for mouse husbandry. We thank L. Berg for NKT tetramer, L. Lanier for anti-NKG2D blocking antibody (CX5), B. Polić for $Nkg2d^{-/-}$ mice and M. Exley for $Cd1d^{-/-}$ mice. This work was supported by National Institutes of Health (NIH) training grant AI07349 (S.N.W.) and research grants AI-17672, AI-081675, CA34461 (R.M.W.), AI46578 (L.K.S.), a German Research Foundation fellowship CO310-2/1 (M.C.) and an institutional Diabetes Endocrinology Research Center (DERC) grant DK52530. The views expressed are those of the authors and do not necessarily express the views of the NIH.

**Author Contributions** S.N.W. designed the study, performed experiments, analysed data and wrote the manuscript; R.M.W. designed the study, analysed data and wrote the manuscript; M.C. and L.K.S. were involved in study design, discussed results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.M.W. (Raymond.Welsh@umassmed.edu).



# AMMS

Academy of Military Medical Sciences

PART 2

**Celebrating 60 years of research at one of China's leading organizations for medical science**

超越
*"Surpassing excellence"*
Erected 2010

**www.nature.com/nature/supplements/amms**
**Email: contact_office@bmi.ac.cn**

## Advancing life science and biotechnology research

The Academy of Military Medical Sciences (AMMS) is one of China's pre-eminent centres for cutting-edge life sciences research and development. Since its founding 60 years ago, the AMMS has attracted top-tier researchers in biochemistry, molecular biology and biotechnology and established itself as a major contributor to advances in public health not only in China but also on the global stage.

Part two of the AMMS supplementary feature explores recent advances in the new fields of proteomics and metabolomics, which are enabling researchers to delve deeper than ever before into the relationships between proteins, genes and metabolites, and providing new targets for drug discovery. Taking a leading role in the international Human Liver Proteome Project, the AMMS has identified the largest human organ proteome expression profile to date, revealing new insights that are critical to the development of novel methods for early detection and treatment of liver diseases.

At a time when China and much of the world faces a rapidly aging population, age-related pathologies such as osteoporosis are of increasing concern. The AMMS is taking great strides in osteoporosis research by developing innovative anabolic treatments to restore bone loss, helping to improve quality of life for sufferers worldwide.

The AMMS' world-class research into developmental biology, RNA bioinformatics and biochips are also highlighted in this issue. The academy currently conducts projects ranging from advanced microRNA research to the application of silicon biochips, and is making pioneering contributions to facilitating data sharing with the wider scientific community. With its successive achievements in life sciences research, the AMMS continues to advance transformative innovations in public health around the globe.

**AMMS RESEARCH**

# A new era in proteomics research

Proteomics is rapidly emerging as a powerful technique in biomedical research for identifying and characterizing proteins involved in key biological pathways, and holds great promise in developing novel diagnostic tools and therapeutic strategies. The AMMS is at the forefront of advancing proteomic health care technologies in China, and is also spearheading large-scale proteomic research initiatives worldwide.

Over the last decade, international efforts to build a comprehensive catalogue of human proteins — a proteome — have greatly accelerated our understanding of the underlying basis of human health and disease. Researchers at the AMMS are taking a leading role in the burgeoning field of proteomics, as evidenced by the establishment of the China Human Proteome Organization in 2002, and the successful implementation of the China Human Liver Proteome Project, involving collaboration with thousands of top scientists and medical experts from more than 50 institutions in China.

At a national level, the strength of China's research capabilities was reinforced by the establishment of the Beijing Proteome Research Center in 2003, and the subsequent launch of Proteome$^{SKY}$, the first and only national laboratory dedicated to proteomics in China. The AMMS currently leads the international Human Liver Proteome Project (HLPP) under the directorship of AMMS president Fuchu He to build on the success of the China Human Liver Proteome Project. To support large-scale proteomic analyses, the AMMS has developed a range of novel technologies including tools and software to enable researchers to perform quantitative studies of specific protein expression and modification, and an advanced bioinformatics platform for data processing, evaluation, mining and sharing. The AMMS was among the top six laboratories out of 27 participating groups to achieve a 100% score in the Human Proteome Organization's Test Samples Study, aimed at establishing global standards in proteomics.

## A world leader in liver research

Liver diseases are a major focus of research at the AMMS, and the academy has been able to provide unparalleled technical support to proteome researchers investigating all aspects of the major physiological and pathological processes occurring in the human liver.

The AMMS' Beijing Proteome Research Center has played a crucial role in defining a set of key scientific objectives that have been adopted by the HLPP. These objectives include characterization of the proteome of the healthy human liver based on expression and modification profiles, development of a protein linkage map and a proteome localization map, and the compilation of sample libraries, antibody libraries and databases to aid research into both the physiological and pathological functions of the liver. Due to the insights offered by the China Human Liver Proteome Project, which involved the systematic analysis of human fetal liver and adult healthy liver, the AMMS has identified the largest human organ proteome expression profile to date, encompassing 6,788 high-quality human liver proteins, of which 3,721 proteins were identified for the first time in the human liver. In AMMS' proteomic studies relating to major liver diseases, researchers have identified proteins associated with fatty liver, viral hepatitis and hepatocellular carcinogenesis, as well as novel biomarkers for early detection of liver diseases and potential therapeutic targets.

## State-of-the-art research facilities

Researchers at the AMMS have used state-of-the-art yeast two-hybrid screening approaches to discover novel protein–protein interactions. Following the selection of more than 5,000 functional molecules in the human liver, a library of 9,600 yeast two-hybrid plasmids has been screened to generate more than 3,400 pairs of protein–protein interaction datasets. The AMMS' high-throughput screening technology is based on a microarray system adapted for 96-well assay plates, enabling researchers to test up to 384 × 384 protein-protein interactions each day.

Fuchu He

Quantitative proteomics research has also benefited from new nanobiotechnologies developed by the AMMS. Researchers have developed a proprietary metal-chelated magnetic nanobead technology that provides higher efficiency for polypeptide enrichment, as well as a non-magnetic nanobead serum treatment system for improved compatibility with mass spectrometry. Using these technologies, serum peptidome profiling has been conducted on 1,000 healthy people and over 2,000 patients with lung cancer, liver cancer, breast cancer, colorectal cancer, prostate cancer and leukaemia, and researchers have obtained high-specificity serum peptidome profiles of lung cancer and non-small cell lung carcinoma, which may aid in the identification of novel disease-associated biomarkers.

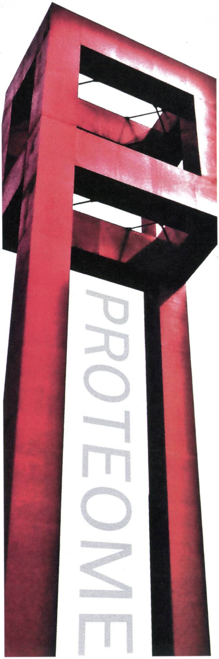
Advances in comparative proteomic technologies at the AMMS have yielded fresh insights into the virulence of pathogenic microorganisms. Analysis of *Shigella*, a representative Gram-negative pathogen closely related to the common human pathogens *Escherichia coli* and *Salmonella*, has led to the discovery of *argT*, a new anti-virulence gene. Screenings of *Streptococcus suis,* a representative Gram-positive pathogen, revealed 32 kinds of immunogenic proteins including 22 novel *S. suis* immunogenic proteins. The AMMS has begun proteomic analysis of Bifidobacteria, the most common type of microbes used as probiotics, and has constructed a proteomic map of *Bifidobacterium longum* NCC2705. AMMS researchers have been able to study the underlying mechanisms involved in fructose transfer and assimilation by obtaining protein expression profiles of *B. longum* grown in different carbon source concentrations.

## Future plans

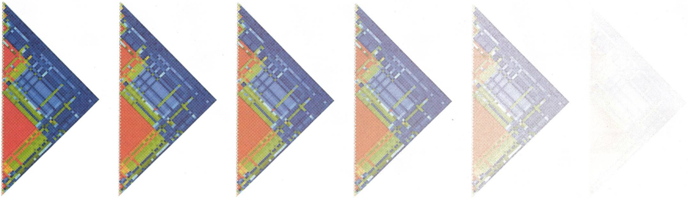
As researchers around the world continue to join forces in characterizing the products of the 21,000-odd protein-coding genes in the human genome as part of the international Human Proteome Project (HUPO), proteomics research is still very much in the ascendancy.

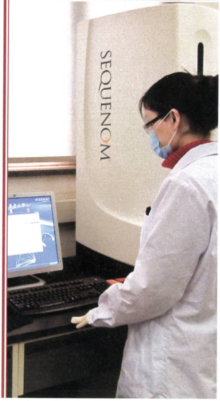
Within the next decade, the AMMS aims to achieve comprehensive coverage of the human proteome and maximize the amount of high-quality quantitative information that can be obtained. These goals also include plans to identify 10,000 protein–protein interactions and achieve quantitative determination of post-translational modifications such as ubquitination, acetylation and methylation, with the overarching aim of understanding the regulation of essential gene expression — and in turn, metabolic diseases — in greater depth.

The AMMS is actively engaged in developing the China Human Proteome Project (CHPP) and is currently laying the groundwork for a newly approved national proteomics co-facility named the 'Pilot Hub of Encyclopedical Proteomix' (PHOENIX) in Beijing's Zhongguancun Life Science Park. With construction due to begin in 2011, the PHOENIX centre will host state-of-the-art facilities for mass spectrometry, optical molecular imaging, cryo-electron microscopy and nuclear magnetic resonance specifically for protein research.



> *The AMMS has identified the largest human organ proteome expression profile to date, encompassing 6,788 high-quality human liver proteins, of which over 3,000 proteins were identified for the first time in the human liver.*





**AMMS RESEARCH**

# Decoding the liver proteome

Researchers at the AMMS Beijing Proteome Research Center are undertaking the formidable task of identifying the entire set of proteins expressed in the human liver, with the aim of both uncovering the proteomic basis of liver development and finding new approaches to diagnose, prevent and treat liver diseases that afflict millions of patients worldwide.

Liver diseases are a major health concern in China. According to estimates by the World Health Organization, more than 93 million people in the country are carriers of hepatitis B virus (HBV) — an infectious illness that causes inflammation in the liver. People with chronic hepatitis B are at risk of developing liver failure, cirrhosis and hepatocellular carcinoma, the most common type of liver cancer. These medical conditions are difficult to cure and costly to manage.

The Beijing Proteome Research Center, one of the AMMS' most prominent research centres, is engaged in an ambitious international project that promises to revolutionize the diagnosis, prevention and treatment of such liver diseases. The Human Liver Proteome Project (HLPP) aims to identify the entire set of proteins expressed in the human liver, as well as functional interactions between these proteins. The results will provide scientists an unprecedented insight into molecular workings of the liver, as well as pathological mechanisms underlying liver diseases.

Xiaohong Qian, Xiaoming Yang and Gangqiao Zhou, leading researchers at the Beijing Proteome Research Center, are major contributors to the HLPP, each being at the forefront of proteomic research and development in their respective fields. Qian is based at the Laboratory of Protein Expression Profiling, where she develops novel techniques for large-scale protein identification and quantitative protein analysis. Yang, based at the Laboratory of Protein–Protein Interaction, studies the functional interactions between liver proteins. Zhou, based at the Laboratory of Single-Nucleotide Polymorphisms, identifies genetic variations in common diseases through genome-wide association studies on Chinese populations.

BECKMAN COULTER

"There are a number of pressing questions about the human liver that we aim to answer," says Qian. "For example, what proteins are expressed in the human liver? Where in the cell are these proteins expressed? What are their expression levels and post-translational modifications — for example, methylation, phosphorylation and ubiquitination?"

Given the sheer number of proteins that might exist in the liver, answering these questions can be a daunting task. As a reference, more than 21,000 protein-encoding genes have been identified in the human genome. The human proteome, however, has the potential to be much larger than the human genome because each gene might encode more than one protein. In addition, a protein may undergo post-translational modifications and its expression level in the liver may vary throughout the day. All these factors add to the complexity of understanding the liver proteome.

Using chemical or affinity-based techniques, researchers at the AMMS Beijing Proteome Research Center have so far identified more than 10,000 representative proteins and their organellar localizations in the liver.

"In the next five years, we would like to identify all the proteins in the liver and obtain quantitative information, such as expression level, for at least half of these proteins," explains Qian.

The Human Liver Proteome Project also aims to uncover the functional interactions between the liver proteins. "We want to identify 'leader' proteins that direct the function of other proteins," says Yang. "However, studying the functional interactions between proteins can be as complex as studying the social interactions between humans."

Using a yeast two-hybrid system to investigate the complex ways in which proteins bind to each other, Yang's research group has so far screened 6,000 liver proteins and identified 3,000 pairs of protein–protein interactions.

Yang's study has cast new light on the essential role of enzymes in the human liver. Some of these enzymes interact not only with proteins involved in metabolism, but also with proteins involved in apoptosis — the programmed cell death that cancer cells have learned to avoid. These enzymes might play an important role in the development of HBV-related hepatocellular carcinoma. In China and the rest of the world, hepatocellular carcinoma remains the most common cause of liver cancer and often follows chronic infection with HBV. However, not all HBV sufferers will go on to develop hepatocellular carcinoma, underlining the complex nature of this particular disease.

Ultimately, researchers aim to use the human liver proteome as a guide to whether a person is at risk of liver diseases. Zhou, who studies common diseases in the Chinese population, is optimistic about the implications. Last year, he initiated the first genome-wide association study aimed at identifying genes that could increase susceptibility towards the development of HBV-related hepatocellular carcinoma in Chinese populations.

"We hope to find proteins that can indicate if a patient is at risk of liver cirrhosis or cancer," says Zhou. "We have used several hundred clinical samples and verified their expression levels, and the results are positive."

> *"In the next five years, we would like to identify all the proteins in the liver and obtain quantitative information, such as expression level, for at least half of these proteins."*

In the groundbreaking study, the genetic profiles of 355 individuals from Fusui County in the south Chinese province of Guangxi were compared with a group of 360 healthy individuals, which led to the identification of 45 genetic variations associated with the incidence of liver cancer. Further investigation identified a single variation associated with the gene *KIFIB*. The results suggested that the influence of this gene variation on the development of liver cancer was eight times greater than the risk arising from other hereditary factors, and indicate that *KIFIB* may be the real suppressor of liver cancer.

Such studies are paving the way to an unprecedented degree of precision and control in disease research. "Our technique is unique in that it allows us to tell you the expression level of each protein," adds Qian. "This means that if you have hyperlipidemia, we can tell you what proteins cause the condition and whether your protein expression levels are out of the norm. This will have a big impact on disease diagnosis."

Since it was first proposed in April 2002, the Human Liver Proteome Project has generated a wealth of data, and as the first initiative on human tissues/organs launched by the Human Proteome Organization (HUPO), aims to develop standard operating procedures for other HUPO initiatives such as the Human Brain Proteome Project, the Plasma Proteome Project, Kidney and Urine Initiative, and Proteome Biology of Stem Cells Initiative.

Decoding the liver proteome is also of paramount interest to the pharmaceutical industry and all those involved in drug discovery and development. "The identification of new drug targets is pivotal to the development of new drugs. If a particular disease is caused by multiple genes, and a gene encodes multiple proteins, then the number of drug targets involved could run into the thousands," says Qian. "The liver proteome can help us identify all candidate drug targets in the liver."

**AMMS RESEARCH**

# Tackling the problems of aging

Global life expectancy is rising rapidly and is imposing new challenges on healthcare systems around the world. With people over 60 years of age set to exceed 20% of the global population by 2050, the search for healthcare solutions for an aging population is ever more pressing. In the post-genome era, attention is turning to the study of genetics and functional genomics for clues to the treatment of lifestyle and age-related conditions. Researchers at the AMMS Beijing Proteome Research Center are at the forefront of China's efforts to tackle these problems with its cutting-edge program of research in areas such as osteoporosis, DNA repair and cancer.

China is the world's most populous country and its rapidly aging population is fast becoming a major public health concern. The sixth national population statistics published in 2010 estimated the country's population to be a staggering 1.34 billion, of which more than 13% were aged 60 or over — a total of approximately 178 million people. One common health problem experienced by elderly citizens is osteoporosis — a disease that results in the reduction of bone mass and bone mineral density leading to weak, fragile bones and increased risk of fracture in patients. In addition to its societal effect, the condition has a significant financial impact and it is estimated that China spends around US$ 2.1 billion on treating osteoporosis every year.

"Osteoporosis is one of the biggest threats in life as it predominantly affects not only elderly citizens, but other people including pregnant women and those with sedentary lifestyles." says Lingqiang Zhang, a principal investigator at the BPRC's Laboratory of Functional Genomics who studies the process of protein ubiquitination and its physiological role in the development of diseases.

A graduate of Yantai and Shandong universities, Zhang joined the AMMS in 2000 as a doctoral student of AMMS president Fuchu He and was rapidly promoted to associate professorship in 2005 and full professorship in 2008. As a scientist born and educated in China, Zhang feels a strong responsibility to develop new approaches to the treatment of osteoporosis for the nation, and the rest of society. The approach taken by Zhang and co-workers at the AMMS involves not only trying to halt the progress of the disease but actually to reverse it. "Current treatments for osteoporosis are relatively simple and focus on preventing bone loss," explains Zhang. "We are trying to find drugs that could restore bone."

Research in Zhang's team at the AMMS focuses on the role of a protein discovered by the team in 2008 called CKIP-1 which specifically suppresses bone formation in mice. Mice lacking the gene governing the production of CKIP-1 have increased bone mass, faster bone formation and improved bone mineral density. Furthermore, AMMS researchers have found that CKIP-1 is vital for the catalytic activity of SMURF, an enzyme previously implicated in bone formation, opening up the possibility of new targets for treatment. Zhang is also investigating the use of biomolecules known as small interfering RNAs (siRNAs) which moderate the expression of a specific gene, to interfere with the expression of CKIP-1. Initial results in rats with induced osteoporosis are promising and early indications suggest that the treatment also works to strengthen the bones of healthy rats.

## Repairing DNA damage

The genetic information contained in the DNA sequence of a cell is vital to the development and functioning of the cell throughout all stages of its life, and damage to the DNA can lead to serious cell malfunction and even cell death. The AMMS carries out research into the mechanisms by which cells protect themselves against premature cell death caused by damaged DNA. One approach targets the identification of key enzymes and attendant co-factors that govern the progress of the cell cycle. In this, AMMS researchers have discovered that a protein involved in bone metabolism is instrumental in

ISTOCKPHOTO/CHRISFARRUGIA

maintaining high levels of an enzyme known as CHK1 which can reduce the susceptibility of DNA to damage by ionizing radiation of some anti-cancer drugs. It is anticipated that these results could lead to the development of sensitizing agents for radiotherapy and chemotherapy.

Researchers from the AMMS are also studying mechanisms regulating gene repair and stabilization of the cell genome. One of the key targets for investigation in this area is p53, a tumour suppressor protein which regulates the cell cycle, along with the TP53 gene which encodes it. Protein p53 acts by helping to conserve gene stability, thereby preventing mutation of the genome, but is also involved in stimulating programmed cell death or apoptosis. However, in over half the known forms of cancer, p53 is mutated and it is further believed that in the remainder mutation of p53, regulators may be responsible for the development of the cancer. Using a multi-strand approach, AMMS scientists identified a protein called APAK that suppresses programmed cell death mediated by p53, but does not halt progression of the cell cycle. The team are now in the process of developing genetic tools, known as RNAis, which can interfere with the production of APAK and can accordingly stimulate apoptosis in cancerous cells. In order to increase understanding of the complex mechanisms surrounding gene repair, the team are also focussing on a suite of biomolecules that are believed to influence the stability of p53. These include SOX4 a protein implicated in the regulation of embryonic development, tumorigenesis, bone development pathways and cell apoptosis, enzymes SMURF1 and SMURF2 that are involved in the process of bone development and PACT, a protein that negatively regulates p53 and which plays an important role in cell growth and embryonic development. By adopting a multi-target approach, AMMS researchers aim to generate a range of genetic resources that will drive the search for new cancer treatments.

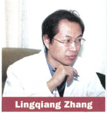
### The genetic basis of cancer

Research into the role of p53 in cell death is just one area in which the AMMS is working to combat cancer, a worldwide health scourge that afflicts many millions every year. The risk of developing cancer depends on a range of factors including lifestyle, environment, and genetic predisposition, and in general increases with an individual's age. Although the first two factors have been studied at length, elucidating and understanding the gene-related causes of cancer is a high-priority challenge in the field of genomic medicine. Building on its program of advanced research into genomics and proteomics, the AMMS is establishing a reputation as a leading centre of cancer genetics research for a range of carcinomas including liver, prostate and breast cancers.

Research into the genetic causes of liver cancer is an area of particular expertise at the AMMS. The academy recently carried out the first ever large-scale genome-wide association study of liver cancer patients which provided further compelling evidence for the role of genetic predisposition for cancer and identified new susceptibility gene regions for the disease (see 'Decoding the liver proteome'). In tandem with research into liver cancer, scientists at the AMMS are also studying the fundamental function of gene pathways and gene products implicated in the formation of other carcinomas, in particular breast cancer and prostate cancer.

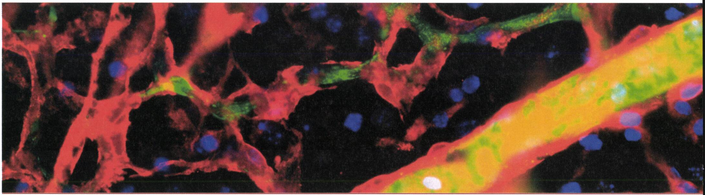
One particularly fruitful line of enquiry involves the investigation of the role of a class of enzymes known as telomerases in the longevity of cells. In healthy cells, the end section of the DNA strands is protected from deterioration that occurs during cell division by a cap known as a telomere, which is partially consumed with each cell division and subsequently rebuilt by the action of an enzyme known as a telomerase. Over many replication cycles however, the ends of the telomere do become shorter until they can no longer protect the termini of the cellular DNA, which causes the cell's reproduction ability to be adversely affected, leading to the cell's eventual death. Whilst this cycle is essential for the correct functioning of healthy cells, malfunction of the system in cancer cells leads to their "immortalization" and corresponding unchecked growth resulting in the formation of a tumour. Researchers from the AMMS have identified a protein, PinX1, that is ascribed with good anti-tumour activity which prevents the immortalization of tumour cells by the inhibition of telomerase.

The telomerase study, along with other cutting-edge research projects carried out by AMMS scientists in the field of fundamental genetics and genomic medicine, is providing important insights into the underlying mechanisms of tumour establishment and growth, which hold great promise for the development of new treatments for cancer that will be of great benefit to patients not only in China but also the rest of the world.



**Lingqiang Zhang**

> *"Osteoporosis is one of the biggest threats in life as it affects not only elderly citizens, but other people including pregnant women and those with sedentary lifestyles."*



AMMS RESEARCH

# Exploring disease through cutting-edge biomedical research

The advent of modern molecular biology methodologies has led to unprecedented insights into the development, onset and treatment of major human diseases. Researchers at the AMMS are investigating the fast-growing field of metabolomics — widely considered to be one of the most promising methods to uncover the molecular basis of human health and disease — as well as demonstrating the outstanding versatility offered by the very latest advances in developmental genetics and RNA-related technologies.

Monitoring metabolism in the human body to gauge the health of an individual has existed since ancient times, but modern-day studies of metabolic pathways — or metabolomics — have come to the fore only in the last few years. As a close cousin of genomics and proteomics, metabolomics offers a unique opportunity to examine genotype-phenotype relationships and uncover the roles of enzymes in key biological processes. Metabolomic profiling is widely regarded as a major step towards a new era of personalized medicine. Technological advancements in liquid chromatography-mass spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR) spectroscopy have played an integral role in the development of metabolomics research, and international research efforts making use of these technologies led to the release of the first comprehensive map of the human metabolome in 2007.

## Advances in metabolomics

The AMMS has established an effective functional metabolomics technology platform based on NMR, GC-MS and LC-MS techniques and data analysis systems, and has conducted extensive studies to evaluate the efficacy and safety of new drugs, as well as the pharmacological mechanisms of Chinese medicinal herbs. AMMS researchers are developing integrated approaches based on metabolomic analyses and animal disease models such as fast-aging mice, and are conducting high-throughput screening of potential diagnostic markers.

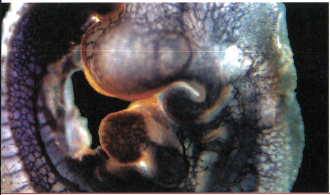
As metabolomics offers a more accurate way of quantifying metabolites than ever before, it is particularly relevant to toxicological studies, and is playing an increasingly valuable role in the early stages of drug development and drug safety assessments. In a toxicological evaluation of Z24, a novel anti-cancer drug candidate, metabolomic analyses conducted at the AMMS have revealed changes in energy expenditure and lipid metabolism in rats. Similar findings have been made in metabolomic studies of Bay 41-4109, an anti-hepatitis B virus compound, leading to new insights into the molecular basis of chemical toxicity.

The mechanisms underlying the efficacy of Chinese herbal remedies continue to be a subject of perennial interest. Metabolomics offers a powerful, systematic approach for the screening and scientific evaluation of Chinese medicinal herbs. Metabolomic analyses of the lateral roots of Carmichael's Monkshood (*Aconitum carmichaelii Debeaux*), are providing a greater understanding of the relationship between alkaloid toxicity and body fluid metabolism in rats. Further developments in metabolomics technologies may play an increasingly active role in the modernization of traditional Chinese medicine.

## Developmental genetics

The deciphering of the human genome at the dawn of the twenty-first century has led to exciting new discoveries in developmental genetics. At the same time, developments in the use of animal models such as knockout mice have been instrumental in revealing the functions of genes that are central to the advancement of immunology and cancer genetics.

At the AMMS' Beijing Proteome Research Center, researchers at the Laboratory of Animal Models are generating many different types of



knockout mice useful for studying signaling pathways and modeling human diseases including cancer, osteoporosis and heart disease.

"Our laboratory generates both transgenic and knockout mice," says Xiao Yang, a principal investigator at the BPRC's Lab of Animal Models. "To generate transgenic mice, we inject foreign genes into fertilized mouse eggs. To generate knockout mice, we inject blastocysts with targeted ES cells."

Both transgenic and knockout mice are important *in vivo* tools for characterizing gene functions, but the questions they help to answer are quite different. Transgenic mice carry foreign genes in their genome and therefore are for use in 'gain of function' experiments. They offer scientists an opportunity to explore potential mechanisms underlying a human disease. In contrast, knockout mice have some of their genes "switched off" and therefore are for use in 'loss of function' experiments. They help scientists identify the role of a specific gene in human diseases.

Adding to the complexity is the fact that mice have a list of "housekeeping" genes that are vital to maintaining basic cellular functions, and inactivating these vital genes could be lethal to mice. Yang's laboratory is special in that they could generate tissue-specific knockout mice — mice that survive longer because the vital gene is inactivated in only one specific type of cells.

"We were the first to use tissue-specific gene knockout technology in China," says Yang, who returned to China in 1999 to open her own mouse laboratory. "Even today, there are fewer than ten laboratories in the country that can generate conditional gene-targeted mice."

So far, Yang's laboratory has generated more than 50 different kinds of transgenic or tissue-specific knockout mice for use in experiments. Most of these experiments revolve around the physiological function of tumour suppressor Smad4 and its mediated transforming growth factor beta (TGF-β) signaling pathway — a pathway that regulates apoptosis, cell cycle and immunity.

"Using our conditional knockout mice, we showed that knocking out Smad4 in keratinocytes (skin cells) can lead to squamous cell carcinoma (skin cancer)," says Yang. "Recently, we showed that knocking out Smad4 in brain endothelial cells can lead to neonatal brain haemorrhage. The results reveal new functions of Smad4-mediated TGF-β signaling in maintaining tissue homeostasis and suppressing cancers and cardiovascular diseases."

The potential applications of transgenic and tissue-specific knockout mice appear to be endless. Apart from molecular genetics and disease models, scientists could use the mice to study microRNA.

"We are now using transgenic mice to study the physiological functions of microRNA — small, non-coding RNA — regulated by TGF-β," says Yang. "We have previously shown that knocking out Smad4 in cardiomyocytes (heart muscle cells) can lead to cardiac hypertrophy (a thickening of the heart muscle), which often results in heart failure. Recently, we have found that some microRNAs, such as miR-24 and miR-27b, can be downregulated by TGF-β and promote the skeletal muscle formation and cardiac hypertrophy, thus validating miR-27b as an efficient therapeutic target for cardiac diseases."

**Xiao Yang**

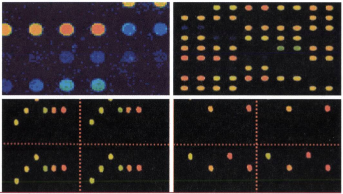
> *"We were the first to use tissue-specific gene knockout technology in China. Fewer than ten laboratories in the country can generate conditional gene-targeted mice."*

### Leading-edge RNA research

Using advanced microRNA (miRNA) data analysis methods, AMMS researchers are investigating novel associations between noncoding RNAs and human diseases, and have identified highly expressed miRNA sequence data from the aging mouse liver and human fetal liver tissue. Additionally, bioinformatics and computational biology have emerged as indispensible tools to study the fundamental behaviour and structure of RNA. The AMMS is actively engaged in developing efficient algorithms and software for improved RNA higher order structure prediction. The academy is leading the way in server technology with the development of the world's first RNA Deleterious Mutation Analysis Server, which can evaluate structural changes as a result of single nucelotide substitutions and an analysis software package called Evolution of RNA Structural Robustness to explore RNA folding landscapes.

The AMMS is also making its computational tools available to the wider scientific community. The academy's small RNA target prediction server has so far provided more than 2,000 services to researchers in over 20 countries. The academy has also developed a database called sRNATarBase, which aims to support large-scale, in-depth research into sRNA function.

In May 2003, the academy completed a preliminary study using RNA interference (RNAi) techniques to combat the deadly SARS virus, and licensed a patent in 2005. The AMMS is one of the first institutions in China to carry out research based on the application of SELEX (systematic evolution of ligands by exponential enrichment) technology, and has demonstrated the potential of cell-SELEX in cancer research. The AMMS is steadily investigating the beneficial effects of a broad range of RNA technologies in the treatment and management of the world's major diseases.



**AMMS RESEARCH**

# Frontiers of biochip technology

The biochip is at the forefront of the rapidly expanding fields of genomics and proteomics, and is instrumental in the development of new approaches to clarifying the complex biochemical processes in cells and improving the diagnosis and treatment of disease. Biochip development underpins many fields of research at the AMMS, where the technology has proved key to developing enhanced treatments for viral hepatitis and improved bone marrow transplants with a number of landmark biochip device approvals.

Since its inception in 1998, the AMMS' biochip technology programme has produced a series of world-class biochips for use in viral hepatitis and bone marrow research, and has secured regulatory approval for numerous biochip-based medical devices. Viral hepatitis is epidemic in China and is a pressing public health concern, with approximately 38 million people diagnosed with hepatitis C throughout the country. In their efforts to tackle the challenges of viral hepatitis diagnosis, the AMMS has successfully developed a protein chip for the detection of hepatitis C virus (HCV) antibodies. An important instrument in proteomics, protein chips can detect proteins in biological samples, such as antibodies in blood, and help in the early diagnosis of disease. This HCV-antibody protein chip was approved for use in clinical diagnosis in 2002, and in the same year obtained national certification as a leading biological product.

The AMMS has also developed a biochip that targets drug resistance in the hepatitis B virus (HBV). Virologic breakthrough of HBV can occur in patients undergoing chemotherapy if the patient has a mutation on what is known as the YMDD locus on the HBV reverse transcription gene. The AMMS created a biochip that is able to detect this YMDD mutation, providing not only genetic diagnosis for HBV patients undergoing chemotherapy, but also the capability to monitor the existence of resistance mutations. The gene biochip was approved for clinical use in 2005 and is the first of its kind to be granted such regulatory approval.

Bone marrow or stem cell transplantation is the sole therapeutic treatment in the fight against leukaemia and other potentially fatal blood diseases. In order to match a recipient to a suitable donor, tissue samples from both the donor and recipient are compared through a technique called human leukocyte antigen (HLA) typing. This technique is important in confirming the compatibility of the donor and recipient in order to minimize later problems in the transplantation process, such as rejection of the transplanted cells. The AMMS has created a new genotyping chip that can determine an incompatible match between donor and recipient by detecting up to 57 types of cell mutations. The chip uses 117 oligonucleotide or molecular probes to hone in on cell mutations in donor and recipient tissue samples. This is the first regulatory approved gene chip designed for HLA-typing analysis, and complements a second genotyping chip produced by the academy, the HLA-DRB1 chip, which uses 12 probes to distinguish seven types of cellular mutation. In addition to the certification of HLA-related biochips, the AMMS has received accreditation of a further four clinical biochip products.

Over the course of biochip development, the AMMS has also made key improvements to various other biotechnology products. The AMMS established a new visual inspection platform incorporating nanogold labelling substrate molecules that enables not only visualized test results, but also greatly enhances the sensitivity of detection. This technology in turn paved the way for the improvement of biochip scanners, resulting in a new portable biochip scanner and analysis system. This scanner offers several advantages compared with the conventional confocal laser scanner, including faster scanning speeds in a smaller, lighter package. The portable biochip scanner is also an economical option, costing approximately one-tenth the price of confocal laser scanners.

These major achievements in biochip development showcase the advanced biochip products created by researchers at the AMMS. In the coming years, the academy will continue devoting its energies to developing more world-class approved biochip products, bringing transformative innovations to healthcare worldwide.

# CAREERS

**JOB SECURITY** UK junior faculty members face poor employment outlook **p.401**

**UNITED STATES** Age of researchers at first grant is increasing, says study **p.401**

**NATUREJOBS** For the latest career listings and advice **www.naturejobs.com**

S. STAFFORD

**OUTREACH**

# Field hospitality

*Hosting guests on research trips can give scientists a chance to showcase their work — but it can also cause distractions.*

**BY LUCAS LAURSEN**

Early in his career, Paul Olsen sat in front of a television, expecting to see his own image. He had hosted a television crew on a research expedition to Manicouagan Crater in Canada, where he and his team were investigating the Triassic–Jurassic boundary in the geological record. Olsen, a palaeontologist at Lamont–Doherty Earth Observatory of Columbia University in Palisades, New York, had spent hours explaining and re-explaining for the camera how scientists used the site to reconstruct ancient ecologies. As the opening credits rolled, Olsen wondered how he would come across on the small screen.

But by the end credits, he was still wondering. All that had appeared on screen was a brief flash of Olsen's disembodied hand, as the documentary focused on other researchers from the expedition. In the decades since then, Olsen has hosted at least a dozen teams of journalists, from National Geographic, the BBC and the Canadian Broadcasting Corporation among others, each time devoting days to assisting them and solving logistical challenges on their behalf. He has learned that a journalist's gratitude does not always translate into a starring role — and nor should it. Although some of his colleagues express disappointment when their hours of storytelling don't make the final cut of a documentary, Olsen acknowledges that journalists must be "fairly ruthless", because "they're there to tell a story". Scientist hosts might find that they're not part of that tale.

Hosting guests such as journalists, policy-makers or teachers during research fieldwork is a gamble. In exchange for a shot at publicizing their work or educating a wide audience, scientists must give up an unknowable amount of time and accept the risk — and the added cost — of taking extra people to a remote location. If it goes well, a visiting teacher might help to spread the word, or a policy-maker might be able to make more informed decisions; but scientists may also end up with recalcitrant or shivering visitors who hole up in their rooms instead of getting close to the fieldwork. To make the most of bringing guests to the field, hosts should plan ahead, communicate their goals, try to head off conflicts and expect to exercise patience.

## BE PREPARED

Hosting may give researchers an opportunity to sell their science. "It theoretically helps your career," says Kevin Krajick, a media-relations manager and science writer at Lamont–Doherty, who helps to connect journalists with scientists. "It's a bit of marketing — organizations such as the National Science Foundation love to see media — and it can help educate the general public."

The impact may be more important for general fields of study than for any one researcher's work. "I don't think there are too many benefits directly to the scientist," says Olsen. Yet he continues to host guests owing to a sense of responsibility to members of the public, whose taxes fund most of his work. Jill Baron, an ecologist at the US Geological Survey in Fort Collins, Colorado, says that bringing policy-makers closer to the wilderness helps to persuade them that such places are worth protecting. Visits can also establish useful relationships. "When they need information they are likely to call you," says Baron, who has hosted policy-makers ranging from local air-quality commissioners to US senators.

Communicating science in the field can help to further a cause and even, in rare cases, catalyse research-vindicating policies. In 2007, after decades of visits by Colorado officials to Baron's field sites, a consortium including the US Environmental Protection Agency, Colorado's air-quality commission and the US National Park Service drafted a plan to reduce nitrogen deposition in the Rocky Mountain National Park. Baron says that without such visits, "it's hard for people to actually visualize" environmental threats. She once spent hours explaining to a mayoral staffer from Denver how agricultural fertilizer used on Colorado's plains infiltrated the Rocky Mountains. After walking all day through the Rockies, the staffer stopped and asked her: "'You mean what we do down in Denver makes a difference up here?'" Baron recalls. "You can talk until you're blue ▶

Left, palaeontologist Paul Olsen (kneeling) hosting a camera crew in Morocco. Right, Kevin Krajick, a science writer, (left) on a field trip in Bangladesh.

▶ but until they see it they won't understand," she says.

When David McGee, a climatologist at the Massachusetts Institute of Technology in Cambridge, hosted freelance writer Douglas Fox on a field trip, he expected to teach Fox about the science involved in his work. But McGee was surprised by what he learned from the experience himself. "It improved my ability to talk about what I do," he says. "You get better at breaking out of the jargon and lingo and assumptions about what is important" to outsiders (see *Nature* **468,** 465–467; 2010). McGee took Fox, who is based in San Francisco, California, on a trip to the dry bed of Lake Bonneville, which once covered much of what is now northern Utah. McGee says he was excited that someone was interested in his work and helping to publicize it. The expedition led to a feature story in *High Country News*, a magazine for the western United States (go.nature.com/qiiyqj); McGee now shares the article with prospective graduate students and postdocs.

As useful as these outreach lessons were, they did cause frustrating and time-consuming complications. One of McGee's Utah field sites was on US military property. McGee had arranged permits long before the expedition — for scientists. But when his military contacts discovered that Fox was joining the trip, they decided that the team would need a military supervisor. They also decided that they were willing to send the supervisor for only one day. So Fox missed out on some of the expedition's scientific research and McGee got stuck with lengthy, last-minute negotiations. "Once you've invited this person along and they've got their plane tickets, you end up with some responsibility to help as an intermediary," he says. "It all worked out in the end, but it was a headache."

McGee also learned that journalists are attracted to the little conflicts that scientists are less eager to broadcast. With Fox around all the time, every prickly exchange among the researchers, no matter how routine, was potential story material. "Sometimes with collaborators you're short about someone's difference of opinion," says McGee. "But that's not the thing you want to see in print." As a result, some of his colleagues were not comfortable with Fox's presence, and McGee now knows that he needs to check with other members of his team before taking a journalist along.

## LESSON PLANNING

A proactive approach before the field visit can help to mitigate misunderstandings, says Krajick. "You want to be prepared," he warns scientists before they take visitors on trips. "Be ready to explain your stuff and have a talk over the phone before going on the expedition to make sure you're on the same page." That allows scientists to lay ground rules such as when they will be available for questions; and it gives journalists a chance to explain when in the editorial process they might want to consult the scientist to check facts.

Olsen discusses mutual expectations with all guests ahead of time. He also prepares handouts for teachers and journalists, summarizing the main points of interest of the field site and the objectives of the expedition. Baron and her students and technicians separate groups of policy-makers into different hiking parties, on the basis of their physical abilities. "Don't take staffers or policy-makers or senators, especially senators, anywhere they're going to be cold or thirsty or get altitude sickness," she advises.

It helps for scientists to have some control over who goes into the field with them. "One of the things we have learned in the past is that the researchers need to be able to pick their teacher," says Janet Warburton, an education coordinator at PolarTREC, a programme based in Fairbanks, Alaska, that gives school teachers a hands-on research experience in the polar regions. PolarTREC filters applicants for motivation and the ability to translate their experience into lesson plans; researchers themselves must ensure that the teacher will be able to cope with the challenges of life and work on their field trip. Lack of filtering can cause wasted opportunities: in one unfortunate case, says Warburton, a teacher not up to the physical challenge of the daily walk to a field site remained in the barracks instead of participating in the fieldwork.

Mark Goldner, a teacher at Heath School in Brookline, Massachusetts, learned about research expectations first hand. Goldner was shortlisted for an expedition to the Svalbard archipelago in the Norwegian Arctic last July. Before he could join the trip, he had to pass a thorough interview with the principal investigator — Julie Brigham-Grette, a palaeoclimatologist at the University of Massachusetts in Amherst. "Julie came right out and said, 'Here's some things we expect of you — what do you think?'" Goldner remembers.

> *"Be ready to explain your stuff and make sure you're on the same page."*

As a trip leader, Brigham-Grette always tells potential participants in her field trips about the physical demands of excursions, such as long walks in the cold, and asks them about their ability to handle responsibilities that include blogging about the trip and helping with expedition equipment. She had one teacher who wasn't able to operate his computer or blog effectively on a trip — and because Brigham-Grette considers communication an important part of a visiting teacher's role, she now asks about those skills. "When I'm in the field, I can't help them navigate that," she says.

Meeting in advance can mitigate stress and

T. MARTIN/ARCUS

Tim Martin, a teacher, tags along with palaeoclimatologist Julie Brigham-Grette in Siberia.

distractions once the trip is under way — in the field, it is sometimes hard for scientists to explain the big picture of the research effort because they are so caught up in the mechanical details of data collection. "When teachers are going out in the field, they're seeing only one piece of the process," says Warburton. "The team is usually very stressed collecting data and may not have the energy to explain the big picture in the moment."

For teacher trips, pre-planning should entail some thought about how the experience will benefit the classroom. Before his own expedition to Siberia with Brigham-Grette, Tim Martin, a teacher at Greensboro Day School in North Carolina, discussed with her how his web design and photography skills could help to translate her science into useful lessons for his students. He also improved his understanding of her work.

Since their expedition together, Brigham-Grette and Goldner have stayed in touch so that she can share her group's analyses of the data that they collected. "The whole goal of PolarTREC shouldn't be about me, or my trip to the Arctic, it's really about my students and the outreach that I can do," says Goldner. "That ongoing collaboration is really important."

### ON THE RECORD

Allowing guests, especially journalists, to participate in expeditions may not be the best choice for every expedition. Anything from precarious logistics to bad weather or sensitive politics might cause scientists to postpone, says Olsen. Krajick adds, "If you're not prepared to go on the record all the time, I don't think you should take someone along." Most fieldwork takes place in too intimate a setting to expect much control over gossip or frayed nerves.

Sharing the experience doesn't have to mean relinquishing all control. Olsen has experimented with one way of dealing with the time demanded by television crews: he has decided to not bother trying to conduct real science for the cameras. Mock observations are enough for television journalists, and not having to worry about botching precise measurements makes answering questions easier.

Indeed, hosting often requires flexibility. "You can't plan which days are for discovery and which are photo days," says McGee. He acknowledges that Fox's visit was rather long at four days, but McGee liked the result — a nuanced depiction of his work.

Field guests can be an asset, even beyond outreach: scientists may be able to put visitors to work and make use of their skills. When Brigham-Grette first offered him a place on her trip, Martin was "ready to mop floors". But the team found more useful things for him to do: he had once worked building houses in needy communities, and had construction skills. On one occasion, Brigham-Grette assigned Martin to work on a drill. He didn't falter when the sub-freezing temperatures chilled the drill's fluids and shut it down. Instead, he assumed the role of foreman. "We brought it inside for warmth," he recalls. "But the exhaust was poisonous, so I built an enclosure to channel the exhaust outside."

Goldner happened to know how to drive boats, so Brigham-Grette asked him to ferry scientists and equipment around. And on a remote field site in Morocco, Olsen once asked a television crew with a large budget to transport some of his team members between sites. When it comes to field-trip guests, he says, "their needs and your needs can overlap." ■

---

**Lucas Laursen** *is a freelance journalist based in Madrid, Spain.*

---

**JOB SECURITY**

## British prospects grim

Recruitment and job security for early-career faculty members in the United Kingdom will become more uncertain as government funding for higher education declines, says a trade union. The University and College Union (UCU) in London released a study of higher-education revenues on 5 January. The report says that the proportion of university income accounted for by government funding, estimated to be 31.6% for 2010–11, will fall to 15% by 2014–15. That will probably lead to a reduction in long-term, renewable contracts for junior faculty members, says Stephen Court, the UCU's senior research officer. "The shifting pattern of income from the government will make university employment more precarious," he says.

---

**BIOMEDICAL SCIENCE**

## Jackson Lab to expand

The Jackson Laboratory has completed plans to open a branch in Connecticut. On 5 January, state governor Dannel Malloy finalized an agreement with Jackson, a biomedical research centre based in Bar Harbor, Maine. The lab will initially lease space from the University of Connecticut Health Center in Farmington. Next year, construction will begin on a permanent facility in Farmington that is set to employ 300 scientists and support staff by 2017. Recruitment is already under way in computational biology and systems genomics for the temporary lab, which will have 27 employees, including scientists, by the end of 2012, says Michael Hyde, Jackson's vice-president for advancement.

---

**UNITED STATES**

## Scientists miss their peak

US biomedical scientists rarely earn their first major grants during their optimum innovative years, concludes a study (K. R. W. Matthews *et al. PLoS ONE* **6**, e29738; 2011). In 2008, the average age of a scientist getting a first grant from the US National Institutes of Health (NIH) was 42, the authors found. But researchers who won Nobel prizes in medicine or chemistry between 1980 and 2010 did their pioneering work at an average of 41 years; 78% did so before 51, the average age of NIH investigators now. Part of the problem is that the NIH is risk-averse and unwilling to fund nascent work, argues Kirstin Matthews, lead author of the study and a science and technology policy fellow at Rice University in Houston, Texas.

# naturejobs

| | |
|---|---|
| **European Head Office, London** | naturejobs@nature.com |
| The Macmillan Building, 4 Crinan Street, London NI 9XW, UK<br>Tel: +44 (0) 20 7843 4961 Fax: +44 (0) 20 7843 4996 | |
| **European Manager:**<br>Nils Moeller,+44 (0) 20 7843 4953 | n.moeller@nature.com |
| **Southwest UK/RoW:**<br>Emma Hare, +44 (0) 20 7843 4944 | e.hare@nature.com |
| **Northeast UK/Ireland:**<br>Matthew Ward, +44 (0) 20 7014 4059 | m.ward@nature.com |
| **France/Switzerland/Belgium:**<br>Julie Aramburu, +44 (0) 20 7843 4994 | j.aramburu@nature.com |
| **Scandinavia/Spain/Portugal/Italy:**<br>Megan Ball, +44 (0) 20 7843 4973 | m.ball@nature.com |
| **North Germany/The Netherlands/Eastern Europe:**<br>Kristin Doelling, +44 (0) 20 7843 4970 | k.doelling@nature.com |
| **South Germany/Austria:**<br>Maria Kubalova, +44 (0) 20 7014 4084 | m.kubalova@nature.com |
| **Natureevents:**<br>Matthew Clare, +44 (0) 20 7014 4015 | m.clare@nature.com |
| **Business Development Manager, Middle East:**<br>Jon Giuliani, +44 (0) 20 7843 4922 | j.giuliani@nature.com |
| **USA Head Office, New York** | naturejobs@us.nature.com |
| 75 Varick Street, 9th Floor, New York, NY 10013-1917<br>Tel: +1 800 989 7718 Fax: +1 800 989 7103 | |
| **US Sales Manager:**<br>Ken Finnegan, +1 212 726 9248 | k.finnegan@us.nature.com |
| **Japan Head Office, Tokyo** | |
| Chiyoda Building, 2-37 Ichigayatamachi, Shinjuku-ku, Tokyo 162-0843<br>Tel: +81 3 3267 8751, Fax: +81 3 3267 8746 | |
| **Asia-Pacific Sales & Business Development Manager:**<br>Yuki Fujiwara, +81-3-3267-8765 | y.fujiwara@natureasia.com |
| **Business Development Manager, Greater China/Singapore:**<br>Gloria To, +852 2811 7191 | g.to@natureasia.com |
| **India** | |
| 3a, 4th Floor, DLF Corporate Park, Gurgaon, 122 002, India | |
| **Advertisement Sales:**<br>Vikas Chawla, +91 1242881057 | v.chawla@nature.com |

## Faculty and Core Facilities Manager Positions: A New Institution for Plant Science and Biotechnology, Chinese Academy of Sciences

The Chinese Academy of Sciences (CAS) is launching a new research institution with an initial emphasis on plant stress biology and expected expansion to other important areas of plant science and biotechnology. With joint funding from the Chinese central government, CAS, and the Shanghai municipality and in affiliation with the Shanghai Institutes for Biological Sciences, the new institution aims to become a world's leading center for plant-related research. It will curate innovative, vibrant, and interactive research groups synergized by a battery of state-of-the-art core facilities.

Highly motivated and talented candidates are sought for faculty positions at various levels, including well-established, mid-level and junior principal investigators (PI) and supporting PIs/core facility managers. Candidates of all nationalities who conduct cutting-edge research on any important areas of plant biology and/or have experience in managing genomics, proteomics, bioinformatics/computational biology, or cell imaging core facilities are encouraged to apply.

The supporting PIs will manage relevant core facilities (50%) and conduct collaborative and independent research (50%). All PIs will be provided with competitive packages including set-up and annual operating funds, and research staff members at various levels as well as with attractive salaries and benefits.

All candidates must have postdoctoral experiences and significant research accomplishments in relevant research fields. For the supporting PI positions, preferences will be given to candidates with extensive experiences in managing relevant core facilities. Qualified candidates should send a cover letter, CV, statement of experiences and future research interests, and arrange to have at least three references sent to:

**Mr. Weijun Wang**
**Address: 320 Yue-yang Road, Shanghai, 200031, China**
**E-Mail: wjwang@sibs.ac.cn**
**Tel: 0086-21-54920027**
**Fax 0086-21-54920078**

JP220625RL

---

$u^b$

b
UNIVERSITÄT
BERN

## TransCure Professorship in Pharmaceutical Sciences, University of Bern, Switzerland

The Medical Faculty of the University of Bern is accepting applications for a faculty position at the level of Full Professor or tenure-track Assistant Professor.

Candidates with strong experience and outstanding record of achievement in a research field related to experimental and clinical pharmacology are being sought. The successful candidate will bring an independent and internationally competitive research program in his/her area of expertise. He/she will furthermore participate in the Swiss National Center of Competence in Research called TransCure (see: http://www.transcure.ch). TransCure is a multidisciplinary team of Swiss experts that focuses on the development of remedies for major human diseases such as diabetes, cancer, heart diseases and neurological diseases based on scientific breakthroughs in the field of transporter and ion channel biology. He/she will provide expertise in experimental pharmacology and preclinical drug development such as ADME profiling and support for the development of therapeutic lead compounds emerging from existing TransCure projects. A well-documented record of independent and innovative research is a prerequisite, including significant publications in high impact journals. Specific experience in transport or membrane biology, however, is not required. In addition, he/she will participate in teaching pharmaceutical sciences, medicinal physiology and biochemistry to students of medicine, veterinary medicine and pharmacy. Fluency in English is a necessary requirement.

The University of Bern is an equal opportunity employer. Women are particularly encouraged to apply in accordance with the NCCR's initiative to support women in science.

The application documents including a letter of intent, resume, publications, funding, a short outline of planned research, and names of 3 references should be submitted electronically to the Dean's Office (felizitas.buss@meddek.unibe.ch). The application deadline is March 1, 2012. For more information, please contact the Chairman of the Search Committee, Prof. Peter M. Villiger, Head, Department of Rheumatology, Clinical Immunology & Allergology, Inselspital, CH-3010 Bern, Switzerland. E-mail: peter.villiger@insel.ch.

The Dean's Office, Faculty of Medicine, University of Bern, Murtenstrasse 11, CH-3010 Bern

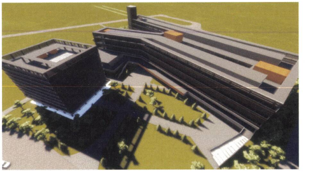
W220671R

---

Visit
**www.naturejobs.com**
to seriously improve your career prospects.

naturejobs
making science work

PHOENIX CENTER



## PHOENIX (Pilot Hub Of ENcyclopedical proteomIX), National Center for Protein Sciences in Beijing, China, is seeking Principle Investigators

PHOENIX has been approved by the Chinese government and is under construction, with the goal of accelerating protein science research and discovery. Academy of Military Medical Sciences (AMMS) is in charge of its construction collaborating with Tsinghua University, Peking University and Institute of Biophysics (IBP) of the CAS. It has two missions, one to perform cutting edge research in protein science and the other to provide services to scientists in the nation and the world. Upon completion, the PHOENIX will be authorized as National Center for Protein Sciences (Beijing). Its headquarter, built by AMMS, is localized in Zhongguancun Life Science Park, and has Beijing Proteome Research Center (BPRC), National Engineering Research Center for Protein Drugs, National Institute of Biological Sciences, Beijing (NIBS), National Engineering Research Center for Biochip (Beijing) as its neighbors. The Headquarter will be housed in a new 37000 square meters modern research building currently under construction, and will be equipped with state-of-the-art instruments in proteomics, imaging, high throughput molecular screening, etc.

We are seeking 20~30 highly motivated scientists to join the staff as principal investigators at all ranks. We welcome applications from all disciplines of protein science, but we are particularly interested in those who study proteomics, protein functions, metabolomics, bioinformatics, imaging and chemical biology. We strongly encourage applications from researchers who are developing and using innovative approaches. To qualify, applicants must hold a Ph.D. or/and M.D. degree, and have demonstrated potential to develop an original, significant research program of international excellence.

We offer competitive salaries and start-up packages. If interested, **send your application package (a cover letter, CV, a detailed research statement, and contact information of three references) to** Dr. Jun Qin (jqin@bcm.edu).

## "Principal Investigators for National New Drug R&D Projects" at the Institute of Pharmacology and Toxicology (IPT), Academy of Military Medical Sciences (AMMS)

IPT, founded in 1958, is one of the leading research institutes in the field of discovery and development of innovative drugs in China. Since its establishment IPT has accomplished the researches for more than fifty-nine marketed drugs and seventy new drug certificates. IPT is also the primary participant in the National Major Project for New Drug Innovation. Ongoing research projects of the institute covers a number of areas, including neurological diseases, cancer, metabolic diseases, cardiovascular diseases, as well as antivirus agents and immunomodulators.

IPT is seeking outstanding scholars and researchers for the position of "Principal Investigator" in all the relevant areas, but not limited to: new drug discovery and screening, chemical synthesis, drug metabolism, bioactivity and safety evaluation, pharmaceutical analysis and formulation. Applicants should be Chinese citizen with a doctoral degree. At least 3-5 years of overseas research or pharmaceutical industry experiences are desired. Successful applicants will be offered with an academic title (professor or associated professor), competitive salaries and benefits. They will have the opportunity to take part in the National Major Project for New Drug Innovation.

Please send nomination and applications to gehuil500@sina.com. All applications should include a CV, a detailed list of publications, a statement of career development plan or research interests, at least three references with the name and contact information.

JP220444R

SLAC
NATIONAL ACCELERATOR LABORATORY

# Laboratory Director of the SLAC National Accelerator Laboratory

## *Operated by Stanford University for the United States Department of Energy*

Stanford University invites nominations and applications for the position of Director of SLAC National Accelerator Laboratory. SLAC, one of the world's leading research laboratories, is a U.S. Department of Energy Office of Science multi-program national laboratory. Established in 1962 at Stanford University in Menlo Park, California, 34 miles south of San Francisco, the laboratory's mission is to explore the ultimate structure of matter and the properties of energy, space and time – at the smallest and largest scales, in the fastest processes and at the highest energies – through robust scientific programs, excellent accelerator-based user facilities and valuable partnerships.

The Director of the Laboratory reports to the President and Provost of Stanford University. The Director is responsible to the President of the University for the effective, responsible and safe operation of the laboratory in support of the mission of the Department of Energy. The Director reports to the Provost for all academic matters involving SLAC and its faculty.
The Director leads the development and implementation of the Laboratory's scientific vision, goals and objectives. The position serves as the lab's highest-level liaison with the DOE and other government, public and private organizations. The Director exercises overall leadership and administration of the Laboratory's programs and operations. The Laboratory maintains an outstanding and growing workforce of approximately 1,600 highly trained and award-winning faculty, scientists, engineers, computer scientists and other professionals, with an annual operating budget of approximately $400 million. Six scientists have been awarded the Nobel Prize for work carried out at SLAC and the future of the laboratory promises to be just as extraordinary. The scientific environment is enriched by over 3,000 visiting scientists annually from universities, laboratories and industrial concerns from the U.S. and foreign countries.

Major existing facilities include the Linac Coherent Light Source (LCLS), the world's most powerful X-ray free electron laser. LCLS creates unique light that can see details to the size of atoms and processes that occur on the femtosecond time scale. Initial LCLS results have met with such success that expansion plans for LCLS-II are already underway. Other facilities include the Stanford Synchrotron Radiation Lightsource based on the 3 GeV SPEAR3 ring; the Instrument Science Operations Center for the Fermi Gamma-ray Space Telescope; the Photon Ultrafast Laser Science and Engineering laboratory (PULSE); FACET, a high-energy beam test facility for accelerator, radiation and material science; and the Kavli Institute for Particle Astrophysics and Cosmology (KIPAC). SLAC will lead the development of the camera for the Large Synoptic Survey Telescope and is an important component of the EXO program to develop increasingly sensitive double-beta decay experiments to elucidate the nature of neutrinos. The Stanford Institute for Materials and Energy Research (SIMES) is a multi-disciplinary joint institute of SLAC Photon Science and Stanford University. The SUNCAT (SUstainable eNergy through CATalysis) Center for Interface Science and Catalysis is a partnership between SLAC and the Department of Chemical Engineering at Stanford University. SLAC maintains a world-leading accelerator research program to ensure that the accelerator-based scientific programs remain at the forefront of the field. The laboratory also has the staff and facilities to advance the state of the art in particle theory, particle detector electronics, microwave power generation and control and scientific computation. In addition, SLAC is in the midst of infrastructure renewal, including major building projects that will transform SLAC into an even more robust laboratory.

Candidates should be recognized scientific leaders with vision, commitment to excellence in research, proven management skills and the interpersonal skills to work effectively with all SLAC constituencies. In addition, candidates should demonstrate the following:

- Strong credentials in scientific research and research management
- Direct experience in developing, constructing and operating major research facilities
- Political acumen and competency in organizational dynamics
- A record of obtaining funding as well as knowledge of how funding decisions are made on the agency – particularly DOE – and Congressional level
- Strategic thinking and execution skills
- Excellent communications skills
- Proficiency in lab operations and community relations

Stanford University is an Affirmative Action Equal Opportunity Employer.

Please note: This position is subject to U.S. Department of Energy approval.

Please submit applications by February 29, 2012 to Director_search@slac.stanford.edu in PDF format **or** hardcopy to US Postal address:

SLAC Director Search
c/o Angel M. Smith
SLAC National Accelerator Laboratory
2575 Sand Hill Road , MS75
Menlo Park, CA 94025

SLAC is operated by Stanford University for the United States Department of Energy.

NW220741R

---

CARL
VON
OSSIETZKY
universität OLDENBURG

The newly established Medical Faculty of the Carl von Ossietzky University of Oldenburg (European Medical School Oldenburg-Groningen, in cooperation with the Rijksuniversiteit Groningen) invites applications for the following tenured faculty position to be filled starting **1 August 2012**:

# Professor (W3) „Human Genetics“

The university is looking for an internationally renowned personality with an outstanding profile in research and teaching. An excellent publication record and experience in one or more of the following fields of human genetics are also required:

- Multifactorial disorders
- Epigenetic processes
- Aging
- Gender and Diversity

Experience with respect to some of the following methods is expected:

- Next-generation sequencing
- Exon-sequencing
- RNA-Seq (whole transcriptome shotgun sequencing)
- genome-wide association studies
- quantitative statistical methods for family-based or population-based genetic analysis

The research areas of the applicant will allow collaboration within the field of hearing research and other neurosensory research fields and be able to contribute to the analysis of patient data in clinical research and thereby provide back-up for health services research. Establishing a special consultation service for the genetics of neurosensory disorders is welcome, provided that an accreditation is possible.

The successful candidate should actively participate in establishing the European Medical School (EMS). The professor is expected to contribute to building up the respective interdisciplinary teaching program of EMS in the field of genetics and to collaborate with the Department of Human Genetics at the University of Groningen.

Active participation in the Centre for Hearing Research (joint institution with the Medical University Hannover) and other structured research institutions at the University of Oldenburg (e.g., the Research Centre for Neurosensory Science) is expected from applicant. The applicant shall contribute to the cluster of excellence „Hearing4all“ which is currently under review and for which a funding decision is expected in June 2012.

General qualifications for the professorial position are specified in §25 NHG.

The Carl von Ossietzky University Oldenburg is an equal opportunity/affirmative action employer. In order to increase the percentage of female faculty members, female candidates with equal qualification will be given preference. Applicants with disabilities will be preferentially considered in case of equal qualification.

Applications including the usual documents (curriculum vitae, list of publications, certificates, diplomas, and list of previous teaching activities) must be postmarked by **1 March 2012** and sent to the **President of the Carl von Ossietzky University of Oldenburg, D-26111 Oldenburg**.

W220676R

---

Medizinische Fakultät Heidelberg

An der Medizinischen Fakultät Heidelberg ist zum 01.03.2013 am Department Innere Medizin die

# W3 Professur für Sport- und Bewegungsmedizin

**(Nachfolge Prof. Peter Bärtsch)**

neu zu besetzen.

Mit der Professur ist die Leitung des Bereichs Sport- und Bewegungsmedizin verbunden. Die/Der künftige Stelleninhaber/-in vertritt das Fach in Forschung, Lehre und Krankenversorgung.

Der/Die künftige Stelleninhaber/-in soll über exzellente Kenntnisse und Kompetenzen in der internistischen Sportmedizin und Leistungsphysiologie verfügen. Die Sportmedizin am Department Innere Medizin bildet zusammen mit Partnern aus dem Universitätsklinikum, dem Deutschen Krebsforschungszentrum, dem Institut für Sport und Sportwissenschaft sowie dem Olympiastützpunkt Rhein-Neckar das Interdisziplinäre Zentrum für Sport- und Bewegungsmedizin (IZSB). Das IZSB dient der Förderung der klinischen und experimentellen Forschung, der interdisziplinären Betreuung von Leistungssportlern und der Optimierung der Lehre für Medizin- und Sportstudenten.

Gesucht wird ein/eine international hervorragend ausgewiesene/-r Forscher/-in auf dem Gebiet der präventiven beziehungsweise rehabilitativen Sport- und Bewegungsmedizin. Erwartet wird die Facharztanerkennung für das Fach „Innere Medizin“ sowie die Zusatzbezeichnung „Sportmedizin“.

Einstellungsvoraussetzungen sind außerdem die Habilitation oder eine gleichwertige Qualifikation, didaktische Eignung sowie Engagement in der Aus- und Weiterbildung. Auf die weiteren in § 47 und § 48, Abs. 3, Satz 4, des baden-württembergischen Landeshochschulgesetzes genannten Einstellungsvoraussetzungen wird hingewiesen.

Als Ersatz für das Liquidationsrecht wird entsprechend dem Beschluss der Kultusministerkonferenz vom 18./19. November 1999 eine gesonderte Vergütungsvereinbarung mit dem Universitätsklinikum getroffen.

Die Universität Heidelberg strebt eine Erhöhung des Anteils der Frauen im wissenschaftlichen Personal an und fordert qualifizierte Frauen nachdrücklich auf, sich zu bewerben. Schwerbehinderte werden bei gleicher Qualifikation bevorzugt berücksichtigt.

Bitte richten Sie Ihre Bewerbung in Papierform mit den notwendigen Unterlagen bis **24.02.2012** an **Prof. Dr. C. R. Bartram, Dekan der Medizinischen Fakultät Heidelberg, Im Neuenheimer Feld 346, 69120 Heidelberg.** Ihre Bewerbungsunterlagen sollten den Kriterien entsprechen, welche Sie unter dekanat@med.uni-heidelberg.de anfordern können.

W220337R

WWW.NIH.GOV

Positions@NIH

THE NATIONAL INSTITUTES OF HEALTH

NIMHD
National Institute on Minority Health
and Health Disparities

**Department of Health and Human Services**
**National Institutes of Health**
**National Institute on Minority Health and Health Disparities**

The National Institute on Minority Health and Health Disparities of the National Institutes of Health is searching for a Clinical Director/Senior Investigator qualified for tenured appointment to direct its clinical research program. Additionally, the Clinical Director is expected to conduct her/his own independent research as a Senior Investigator. As Clinical Director/Senior Investigator, the incumbent is responsible for the leadership of the intramural NIMHD clinical program located at the NIH Clinical Center in Bethesda, Maryland. The Director of the clinical research program/Senior Investigator position is administratively within the Office of the Scientific Director, NIMHD. The successful candidate will oversee a combination of studies, including outpatient, inpatient, epidemiological, clinical protocols, and laboratory studies. Emphasis will be placed upon investigators with a primary research interest in clinical research; however, the selected candidate must adopt the NIMHD strategies in utilizing an interdisciplinary approach toward the science of prevention, diagnosis and treatment of chronic diseases that disproportionately affect the underserved populations in the U.S.

**QUALIFICATIONS:** Candidates must have an established and nationally or internationally recognized record in the conduct and supervision of clinical research in a discipline relevant to the field of minority health and health disparities research. Candidates should also have an established record of experience with clinical training, institutional review boards, and compliance with current clinical safety and regulatory issues. Candidates must have a Doctor of Medicine degree from a school in the United States or Canada accredited by the Liaison Committee on Medical Education (LCME) in the year of the applicant's graduation, or a Doctor of Medicine or equivalent degree from a non-LCME accredited medical school that is certified by the Educational Commission for Foreign Medical Graduates (ECFMG) or hold a Fifth Pathway certificate issued prior to December 31, 2009. Candidates must be board certified or board-eligible in a medical specialty approved by the American Board of Medical Specialties (ABMS), and have the ability to obtain medical licensure in the U.S.

**SALARY/BENEFITS:** The Clinical Director, NIMHD, will be appointed at a salary commensurate with his/her qualifications and experience. Full Federal benefits, including leave, health and life insurance, retirement and savings plan (401K equivalent) will be provided. This position is subject to public financial disclosure requirements.

**HOW TO APPLY:** Interested candidates should send a letter of interest, including a brief description of research and administrative experience; curriculum vitae and bibliography; and full contact information for three to five individuals who may be contacted to provide letters of reference. Application packages should be sent to the **National Institutes of Health (NIH), National Institute on Minority Health and Health Disparities, c/o Dr. Richard Cannon, Chair, Search Committee, 6707 Democracy Blvd., Suite 800, Bethesda, Maryland 20892,** or electronically: keta.lawson@nih.gov; **Subject Line: NIMHD Clinical Director Search.** For further information about the position, please contact Ms. Keta Lawson at 301-594-3986 or keta.lawson@nih.gov. ALL information provided by the candidates will remain confidential and will not be released outside the NIMHD search process without a signed release from candidate. The NIH encourages the application and nomination of qualified women, minorities, and individuals with disabilities.

**APPLICATIONS MUST BE RECEIVED BY MONDAY, MARCH 19, 2012**

**DHHS and NIH**
**are Equal Opportunity Employers**

NW220687R

THE NIH IS DEDICATED TO BUILDING A DIVERSE COMMUNITY IN ITS TRAINING AND EMPLOYMENT PROGRAMS

ecetoc

EUROPEAN CENTRE FOR ECOTOXICOLOGY AND TOXICOLOGY OF CHEMICALS

## ECETOC Secretary General

Founded in 1978, ECETOC is the foremost voice of chemical industry in Europe concerning health and environmental science. The activities of the association are dedicated to the development and promotion of top quality science in risk assessment of chemicals. To this end, a permanent secretariat coordinates the activities of industry experts in collaboration with government and academic scientists in task forces and workshops. The Secretary General is responsible for the execution of the strategy of the association, for the management of the secretariat and the work programmes adopted by the Scientific Committee. He/She is responsible to the board of management for the financial affairs of the association and to the assembly of members for the conduct of the science programme along with its promotion to external stakeholders in government, science and industry.

The ideal candidate would have excellent communication, collaboration and management skills, a higher degree qualification and experience that gives evidence of credibility of scientific accomplishments in areas which are relevant to the risk assessment of chemicals. Experience in a technical or government relations role in industry and/or the regulatory community would be an advantage, as would a working knowledge of the laws and regulations applicable to various types of chemicals in Europe (REACH, Pesticides, Biocides, Pharmaceuticals etc.). The candidate should have an innovative attitude that enables ECETOC to address current and future needs in a complex environment.

The position is based in Brussels and offers an attractive salary, benefits and relocation package. Applications should be accompanied by a C.V. giving full details of career history and be addressed to: Geneviève Gérits, ECETOC AISBL, 6 Avenue Edmond Van Nieuwenhuyse, 1160 Auderghem, Brussels, Belgium.
genevieve.gerits@ecetoc.org.

W220563R

---

hhc
human health care

Eisai

## SCIENTIST, BIOPHARMACOLOGY

**Eisai Ltd. (ESL),** a part of Eisai Co., Ltd, focuses on the development of innovative medicines for the treatment of neurodegenerative diseases.

An opportunity exists for a highly-motivated scientist to join the Neurosciences Product Creation Unit (PCU) based in labs at Eisai's new state of the art European Knowledge Centre facility in Hatfield, UK.

### Neurochemist / Neurophysiologist

We are currently searching for candidates with a PhD degree in neuroscience / neuropharmacology / pharmacology / biochemistry or a related discipline. The candidate should have current, hands on lab-based experience in one or more of the following techniques:

- in vivo microdialysis techniques, including all analytical support functions (i.e. HPLC with ECD and FD) for the measurement of brain neurochemicals;
- in vivo electrophysiology
- electroencephalography (EEG) monitoring.

A thorough practical knowledge of modern technologies and software for the evaluation of data generated via one or more of these methods is required.

All applicants must be enthusiastic, driven and self-motivated research scientists, preferably with a strong neuroscience background. You will be a recognised team player, with excellent written and verbal communication skills and an ability to interact effectively within a highly multidisciplinary environment. Experience within a drug discovery environment would be an advantage but is not essential.

To apply, please send your curriculum vitae, a covering letter and details of two professional referees to: HR Department, Recruitment, Eisai Europe Ltd, European Knowledge Centre, Mosquito Way, Hatfield, Hertfordshire, AL10 9SN or by email to Recruitment@eisai.net.

The deadline for receipt of applications is **17th February 2012**.
*For more information about our company, please visit:* ***www.eisai.co.uk***

U220747R

## More than 100 Professorships & Lectureships

The University of Manchester has a proud record of achievement and an ambitious vision for the future.

The UK's largest university, we're committed to pioneering research, often across disciplines, which also enriches the learning environment for our students.

We have many world-leading scholars on our staff, including three Nobel Prize winners, and we are now looking to appoint more than 100 highly respected academics from around the world to help us to achieve our ambitious goals for research and education.

Here in the heart of a vibrant city, you'll collaborate with brilliant minds to solve tomorrow's challenges today. You'll also educate and inspire the leaders of the future.

If you are at the forefront of your field of expertise, help us to create more world firsts, as we make a significant strategic investment. Visit our website to view our current opportunities and apply online.

# WORLD-LEADING MINDS ANSWERING THE WORLD'S BIGGEST QUESTIONS

ENGINEERING & PHYSICAL SCIENCES | HUMANITIES | LIFE SCIENCES | MEDICAL & HUMAN SCIENCES

MANCHESTER
1824
The University of Manchester

**THE FUTURE, MADE IN MANCHESTER**

**www.manchester.ac.uk/worldleadingminds/nat**

U219522R

# A Jefferiss Research Trust funded initiative in HIV and Infectious Diseases

The Jefferiss Research Trust was set up by the late Dr F.J.G. Jefferiss in 1984 to advance education and research into zoonotic and emerging infectious diseases and sexually transmitted infections for the benefit of the public, and to make the results of the research available to the public in the appropriate manner.

The Trust is keen to support two Ph.D. studentships and two post-doctoral fellowships in these areas of research and is inviting applications now for selection in March 2012, with a flexible/negotiable commencement date in 2012.

For the studentships, applicants should be in possession of or be expected to achieve in 2012 a first class or upper second class degree in microbiology, biochemistry or related discipline. Applications should include a detailed CV, a description of the project **(no more than 2 pages)** that will constitute the proposed Ph.D., the name and CV of the proposed supervisor and a letter from him/her and from the host establishment confirming ability and willingness to accept the studentship and explaining how the project fits into existing research activities. The studentship stipend is £17,500 pa.

For the fellowships, applicants will be expected to produce a coherent and innovative research project **(described in no more than 2 pages)** and work in an established research department with a first class publication record. The CV of the line manager and a letter of support from him/her should be provided, as well as details of the research activities in the host department and an indication of how the work of the Jefferiss Trust Fellow will fit into the existing research programme. A letter of support from both the Head of Department and the host institution should accompany each application. The host department should be well-equipped, as there will be no provision in the grant for equipment or infrastructure costs.

All applicants who hold a Ph.D. and have an exceptional research record in any area of infectious diseases will be given serious consideration. Salary will be, depending on experience, in the range £31,300-£42,160, inclusive of London Allowance.

Interested applicants should send in addition to the aforementioned information, a brief description of current research activities, a statement of future objectives and the name and contact details of three referees to :

The Chairman, The Jefferiss Research Trust, c/o Monro Fisher Wasbrough LLP, Solicitors, 8 Great James Street, London WC1N 3DF email: ddcm@monro-fisher.com

The closing date for applications is the 10th February 2012. Interviewing of short listed candidates will take place in early March 2012 at the above address. There are no formal application forms.

U220129R

---

UNIVERSITÄTSMEDIZIN
GÖTTINGEN
UMG

The Department of Neurorehabilitation Engineering of the University Medical Center Göttingen, Georg-August-University, invites applications for the following two positions:

# Professorship in Neural Signal Processing

(W2, 5 year tenure track)

# Professorship in Neural Interfacing

(W2, 5 year tenure track)

The future professorship holders will be associated with the Department of Neurorehabilitation Engineering. Their research is expected to strengthen the Faculty's scientific focus within the area of neuroscience and neurobionics. The new faculty shall complement existing expertise of the Department of Neurorehabilitation and contribute to the research programs of the ERC Advanced Research Grant DEMOVE ("Decoding the Neural Code of Human Movements for a New Generation of Man-machine Interfaces") and other research grants held by the Department. The Göttingen Research Campus offers a rich scientific environment, where interactions, collaborations and perspectives for additional collaborative funding initiatives are available also beyond the Department of Neurorehabilitation. Of particular relevance are the Bernstein Center for Computational Neuroscience, the Bernstein Focus for Neurotechnology and the Center for Nanoscopy and Molecular Physiology of the Brain.

For the **W2 Professorship in Neural Signal Processing,** research addressing the following topics is needed to complement and complete the existing profile of the Department of Neurorehabilitation Engineering: theoretical aspects of neural signal processing, in particular in the field of time-frequency analysis, multicomponent AM-FM representations for nonstationary signal analysis, blind source separation, spike sorting, and the applications of pattern recognition techniques and system identification frameworks to classify stochastic signals. Applicants should have a strong academic background. Experience within theoretical as well as applied signal processing is required. The applicant should provide evidence of international recognition within the field of neural/biomedical signal processing.

For the **W2 Professorship in Neural Interfacing,** research addressing the following topics is needed to complement and complete the existing profile of the Department of Neurorehabilitation Engineering: nerve and muscle implantable electrode design, experimental expertise in animal models for implantable electrode testing, expertise in bioelectricity and bioelectric modeling, peripheral neural prostheses. Applicants should have a strong academic background. The applicant should provide evidence of international recognition within the field of electrophysiology and neural prostheses, as evidenced for example by the membership in Editorial Boards, conference program committees, invited keynote lectures etc.

More Information about the following link: http://www.med.uni-goettingen.de/de/content/ueberuns/4548.html

Applications please exclusively online via the following link: http://www.med.uni-goettingen.de/de/content/ueberuns/berufungen_onlinebewerbung.html

Questions to the online application please address Berufungen@med.uni-goettingen.de

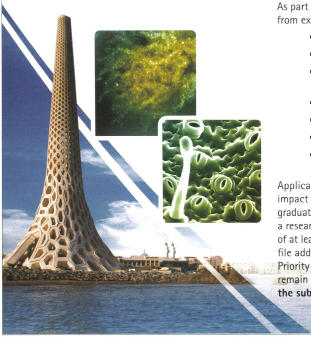
W220490R

# FACULTY POSITIONS: Open Rank

جامعة الملك عبد الله
للعلوم والتقنية
King Abdullah University of
Science and Technology

King Abdullah University of Science and Technology (KAUST) is an international graduate-level, merit-based research university dedicated to advancing science and technology through innovative and collaborative research and to address challenges of regional and global significance.

Located on the Red Sea coast of Saudi Arabia, KAUST offers superb research facilities, generous assured research funding, and internationally competitive salaries. Further information can be found at: *http://www.kaust.edu.sa.*



As part of the expansion of the Bioscience Program we invite applications from exceptional candidates in the areas of:

- Plant biology with an emphasis on abiotic stress responses
- Cell and molecular biology
- Biophysics with an emphasis on structure determination by NMR and EM
- Systems biology
- Pathogen biology
- Genomics and functional genomics
- High resolution imaging applied to biological processes

Applicants should have a proven track record, as well as the ability to establish a high impact research program and a commitment to teaching and research training at the graduate level. Applications must include a complete curriculum vitae with publication list, a research plan, a statement of teaching interests, and the names and contact information of at least three references. Please submit your application electronically as a single PDF file addressed to the Bioscience Search Committee at *BioS_Search@kaust.edu.sa.* Priority will be given to applications received before 29 February 2012 and positions will remain open until filled. **Please identify the position in which you are applying to in the subject heading of the email.**

NW220640R

---

# Contributing to the quality of life

The biobased economy will prove to be essential to provide the world with essential resources for food, chemicals, materials and energy. To boost education and research in the field of 'Biobased Science & Technology', Wageningen University is looking for full professors for the chairs 'Biomass Refinery & Process Dynamics' and 'Biomass Commodity Chemistry'. Both professors will perform new fundamental research leading to technological solutions for the use and processing of biomass and will provide education programs in the particular fields.

## Professor Biomass Refinery & Process Dynamics

The Chair 'Biomass Refinery & Process Dynamics' acts on the cutting edge of technological accessibility of raw biomass and process dynamics. Location: Wageningen, The Netherlands. Vacancy number: WU-2012-HL001.

## Professor Biomass Commodity Chemistry

The chair of 'Biomass Commodity Chemistry' focuses on downstream processing and treatment of macromolecules from biomass resulting in (building blocks of) commodity chemicals for the world market. Location: Wageningen, The Netherlands. Vacancy number: WU-2012-HL002.

### At Wageningen University

Both chairs are part of Wageningen University.

### Interested?

Apply now or view our other vacancies at **www.jobsat.wur.nl**.

*Acquisition regarding this vacancy is not appreciated.*

*Wageningen UR (University and Research centre) is an international leader in education and research. This position is due to the fact that we have been successfully working on the quality of life for many years now. Consequently, we are always looking for talented colleagues. Professionals who want to put their expertise to work for tomorrow's world.*

WAGENINGEN UR
*For quality of life*

**www.jobsat.wur.nl**

W220658R

MDC MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE BERLIN-BUCH
BIMSB THE BERLIN INSTITUTE FOR MEDICAL SYSTEMS BIOLOGY

The MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE (MDC) BERLIN-BUCH and the HUMBOLDT-UNIVERSITÄT ZU BERLIN (HU) invite applications for the following position:

## Associate Professor for Computational Biology

**(Bes.Gr. W2, Reg.-no. 10273/2012)**

The MDC is a biomedical research institute dedicated to interdisciplinary research in the areas of (i) Cardiovascular and Metabolic diseases, (ii) Cancer, and (iii) Function and Dysfunction of Nervous System. The institute is a member of the Helmholtz Association of National Research Centers, supported by the Federal Government of Germany and the Senate of Berlin.

The MDC is expanding in the area of Systems Biology, through the **Berlin Institute for Medical Systems Biology (BIMSB)**, with high-level recruitments and state-of-the-art scientific technology platforms for genomics, transcriptomics, metabolomics and proteomics.

We are seeking applications from outstanding individuals with an international reputation and experience in Systems Biology, who develop and apply computational approaches to understand the biology of gene regulation in eukaryotic organisms. Their aim should be to decipher the complex regulatory networks of transcription, translation or epigenetic regulation by analyzing and modelling data from diverse sources of information (e.g. deep sequencing, mass spectrometry, image analysis). Successful candidates will conduct visionary independent research, obtain extramural funding and engage in the further development of the BIMSB and collaborative projects in Berlin. Knowledge of the German language is not required.

The successful applicant will be jointly appointed by the MDC and the HU. Laboratory space and funds will be provided by the MDC. Candidates are required to fulfill the regulations for a professorship according to § 100 of the "Berliner Hochschulgesetz".

The HU and the MDC seek to increase the proportion of female faculty members. Thus, qualified women are particularly encouraged to apply. Physically disabled persons will be preferred, if they are equally qualified.

For enquiries about the position, please contact Prof. Dr. Nikolaus Rajewsky (rajewsky@mdc-berlin.de). For further information about the MDC/BIMSB and the HU please visit our websites www.mdc-berlin.de and www.hu-berlin.de.

Please submit your electronic application (in one pdf file, maximum size 1 MB) including reference number, curriculum vitae, list of publications, an outline of present and future research plans and other relevant material by March 2, 2012 to Nadine Ewald (**nadine.ewald@mdc-berlin.de**).

Applications should be addressed to:

W220729R

**MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE**
**Prof. Dr. Walter Rosenthal • Scientific Director**
**Robert-Rössle-Str. 10 • 13125 Berlin / Germany**

---

MDC MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE BERLIN-BUCH
BIMSB THE BERLIN INSTITUTE FOR MEDICAL SYSTEMS BIOLOGY

**The HUMBOLDT-UNIVERSITÄT ZU BERLIN (HU) and the MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE (MDC) BERLIN-BUCH invite applications for the following position:**

## Associate Professorship for "Modelling of Molecular Interaction"

(Bes.Gr. W2)

Applicants with an established track record in any aspect of theoretical biophysics, molecular modelling or related fields will be considered, although preference may be given to applicants whose research complements existing research in the department. The candidate is expected to demonstrate a strong commitment to excellence in teaching theoretical physics and mathematical modelling as applied to biological problems in undergraduate and graduate programs for biology and biophysics students.

The professorship will be established at the Institute of Biology at the Faculty of Mathematics and Natural Sciences I of HU and will start at the earliest possible time. The initial appointment is limited to 5 years. Details of the endowment of this position can be enquired from Prof. Dr. Edda Klipp (edda.klipp@rz.hu-berlin.de).

Applicants must meet the legal requirements for appointments of professors in accordance with § 100 of the "Berliner Hochschulgesetz".

HU and MDC seek to increase the proportion of women in research and teaching, and specifically encourage qualified female researchers to apply. Applications of researchers from abroad are welcome. Preference will be given to disabled persons with equal qualifications. Applicants with migration background are highly welcome.

Applications should include a curriculum vitae, list of publications and previous teaching experience, a statement of research and teaching interests, and copies of up to five selected publications and should be sent by referring to reference number **PR/001/12** by March 2, 2012 to the Humboldt-Universität zu Berlin, Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I, Prof. Dr. Herrmann, Unter den Linden 6, 10099 Berlin. Application materials will not be returned. Therefore, you are requested to send only copies of all documents. To accelerate the process, applicants are kindly requested to send their application materials both in written form as well as electronically via https://www2.physik.hu-berlin.de/ssl/modelinter/

W220734R

---

## Neurobiology of Disease Faculty Position

**Gladstone Institutes** and
the University of California, San Francisco

The Gladstone Institute of Neurological Disease and the University of California, San Francisco (UCSF) invite applications for a faculty position at the level of Associate or Full Professor. Of particular interest are neuroscientists and neurologists with broad expertise in chemistry, computational neuroscience, systems biology or neuroimmunology, who are interested in investigating molecular or cellular mechanisms and novel therapeutic strategies relevant to Alzheimer's disease, other neurodegenerative disorders, schizophrenia, autism or related disorders. Primary criteria for appointment will be outstanding records of innovative research and academic performance, including landmark papers in leading journals, as well as proven abilities or high potential to establish a rigorous and substantial independent research program, provide exemplary mentorship, and engage in fruitful collaborations.

The successful candidate will join an interactive group of investigators in Gladstone's state-of-the-art research facility at UCSF's Mission Bay campus. S/he will have joint appointments in the Gladstone Institute of Neurological Disease, the Department of Neurology, the Neuroscience Graduate Program and the Biomedical Sciences Graduate Program at UCSF. Excellent salary support is provided. Women and underrepresented minorities are especially encouraged to apply. The search will continue until the position is filled. However, to ensure full consideration, applications should be received by February 29, 2012.

Please send curriculum vitae, description of achievements and research interests, and the names of three references to:

GIND/UCSF Search Committee
**Gladstone Institute of Neurological Disease**
1650 Owens Street, San Francisco, CA 94158
gindsearch@gladstone.ucsf.edu

GLADSTONE

The J. David Gladstone Institutes and UCSF are Affirmative Action/Equal Opportunity Employers. They undertake affirmative action to assure equal employment opportunity for underutilized minorities and women, for persons with disabilities, and for Vietnam-era veterans and special disabled veterans.

NW220651R

naturejobs
Attention Employers!
Stay up to date with deadlines and upcoming *Naturejobs* features:
Sign up for **FREE** employer e-alerts:
www.naturejobs.com/ealerts

MDC MAX DELBRÜCK CENTER
FOR MOLECULAR MEDICINE
BERLIN-BUCH

BIMSB THE BERLIN INSTITUTE
FOR MEDICAL SYSTEMS BIOLOGY

The MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE (MDC) BERLIN-BUCH is inviting applications for the positions of:

## Independent Junior Group Leader in Systems Biology

**(Reg.-no. 10271/2012)**

The MDC is a biomedical research institute dedicated to interdisciplinary research in the areas of (i) Cardiovascular and Metabolic diseases, (ii) Cancer and (iii) Function and Dysfunction of Nervous System. Additionally the MDC is expanding its impact in the field of Systems Biology and started the **Berlin Institute for Medical Systems Biology (BIMSB)**. The institute is a member of the Helmholtz Association of National Research Centers, supported by the Federal Government of Germany and the Senate of Berlin.

The BIMSB is currently recruiting senior and junior scientists and will move to a new research building in the center of Berlin within the next years.

We are now seeking applications from outstanding individuals working on gene regulation, RNA biology, genetic variability, epigenetics, developmental and/or stem cell biology. Applicants using imaging and/or high-throughput technologies for experimental, computational and theoretical approaches are encouraged to apply. Experiences in highly successful post-doctoral research are required.

Successful candidates will conduct independent research as well as engage in collaborative projects with groups at the BIMSB, MDC and other research institutions in Berlin. We offer Tenure-Track Positions, which include funding for 5 years with possible extensions of two-time 2 years, staff of one postdoc, one PhD student, one technician, as well as administrative support and fixed annual consumables.

Enquiries about the scientific positions and perspectives should be addressed to Prof. Dr. Nikolaus Rajewsky (rajewsky@mdc-berlin.de). For further information, please visit our website www.mdc-berlin.de/en/bimsb

The MDC is an equal opportunity employer committed to excellence through diversity. Applications of women are explicitly encouraged.

Applications should be sent by March 2, 2012, including curriculum vitae, list of publications, outline of present and planned research and other relevant material in one pdf file, maximum size 1 MB, to Nadine Ewald (**nadine.ewald@mdc-berlin.de**).

W220728R

**MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE**
**Robert-Rössle-Str. 10**
**13125 Berlin / Germany**

Zertifikat seit 2009
audit berufundfamilie

---

Obra Social
Fundación "la Caixa"

CNB
CENTRO NACIONAL DE BIOTECNOLOGIA

## The Spanish National Centre for Biotechnology (CNB-CSIC) invites applications for

# 10 PhD Fellowships

## in collaboration with Fundación La Caixa.

Fellowships are available for a period of 4 years, starting in September – October 2012.

The CNB is dedicated to improving human and animal health, as well as solving environmental and agricultural challenges. The excellence of research conducted ranks the CNB among the leading research institutions in the world (as evaluated by EMBO in 2005).

*The major areas of study are:*

- Macromolecular Structures
- Molecular and Cellular Biology
- Microbial Biotechnology
- Plant Molecular Genetics
- Immunology and Oncology
- Systems Biology

**Highly qualified and motivated candidates from all over the world are invited to submit their application before February 17th 2012.**

*See our website www.cnb.csic.es for additional information and application procedure.*

W220767R

---

mfpl
MAX F. PERUTZ LABORATORIES

## VIENNA INTERNATIONAL POST-DOCTORAL TRAINING IN MOLECULAR LIFE SCIENCES

The Max F. Perutz Laboratories offer a unique post-doctoral training program, 'Vienna International Post-Doctoral Training in Molecular Life Sciences' (VIPS), with support from the Austrian Ministry of Science and Research and the City of Vienna.

### Join us in reinventing the postdoc career!

We take the traditional **postdoctoral fellowship** and add **extra independence, funding** and **support.** We place it in our thriving research environment to foster scientific excellence.

**What we offer:**

- A fast-track to the next step in your career
- Support in research design
- Independence: your own research budget and travel funds
- Grant application support
- Mentoring and coaching
- A small, friendly, but nonetheless stimulating scientific environment
- Career guidance and training
- Competitive salary and full health coverage
- A welcome package including assistance with visas
- On-site childcare

**Our profile:**
MFPL is part of the Vienna Biocenter Campus, a center for world class research. We offer a wide range of research areas, from Biochemistry, Cellular Biology, Immunobiology to Organismal Biology, Bioinformatics, Structural Biology and Neurosciences.

**Apply to spend your post-doctoral training in a thriving scientific environment in the world's most liveable capital city!**

Application Deadline 29 February 2012
**www.mfpl.ac.at/vips**

Contact: Gabriele Permoser, e-mail: **vips@mfpl.ac.at**

The Max F. Perutz Laboratories are a joint-venture of the University of Vienna and the Medical University of Vienna.

universität wien MEDICAL UNIVERSITY OF VIENNA BMW_F CityofVienna

W220682R

---

**John Innes Centre, Norwich**

## Group Leader Positions

John Innes Centre

The John Innes Centre is an independent, international centre of excellence in plant science and microbiology (full details on http://www.jic.ac.uk). Our mission is to carry out fundamental research that underpins strategic advances for societal benefit. The scientific research at the Centre makes use of a wide range of disciplines in the biological and chemical sciences, including cell biology, computational and mathematical biology, biochemistry, chemistry, genetics and molecular biology.

We are seeking to recruit at least two new Group Leaders who will run internationally excellent and innovative research programmes focused on plant or microbial systems. We are especially interested in the areas of plant response to the environment, plant cell biology or plant growth and development. One of the appointments will be as part of the Earth and Life Systems Alliance (www.elsa.ac.uk), a collaborative project with the University of East Anglia. For this we are interested in plant biologists linking mechanistic with evolutionary and ecological understanding.

Appointments can be made at any level dependent on the candidate's experience with a tenure-track system for junior faculty. Research support and excellent laboratory facilities are available.

Application forms and further information can be obtained from http://jobs.jic.ac.uk or Human Resources, The Operations Centre, Norwich Bioscience Institutes, Norwich, NR4 7UH, UK, 01603 450462 quoting reference 1002169. Please apply on-line by March 20th 2012 via the JIC web-site including CV, names of three academic referees, statements of current and future research plans.

The John Innes Centre is grant-aided by the Biotechnology and Biological Science Research Council and is an Equal Opportunities Employer.

POSITIVE ABOUT DISABLED PEOPLE

U220735R

UNIVERSITY OF
LIVERPOOL

**Faculty of Health and Life Sciences**
**Institute of Integrative Biology**

## 3 Chairs/Readers/Senior Lecturers

**Job Ref: A-577068/N**

**Chair – Salary Negotiable**
**Reader/Senior Lecturer - £48,247 - £59,304 pa**

## 2 Tenure-Track Fellows (Leading to Lectureships)

**Job Ref: A-577069/N**

**£37,012 – £46,846 pa**

For the Chairs/Readers/Senior Lecturers posts we are seeking individuals with strong track records of international excellence in research and research leadership in your respective fields and previous institutions.

The Tenure-Track Fellows posts will suit early career researchers with evidence of significant research achievements including those currently holding independent funding.

Whilst we will consider applications covering the full breadth of the life sciences we are especially keen to attract candidates in the following key areas:

- Computational Biology and Modelling
- Evolution, Ecology and Behaviour
- Functional and Comparative Genomics
- Cell and Molecular Dynamics, Regulation and Signalling

We are also interested in expanding the following enabling technologies – Cell/Molecular Proteomics; Cell/Molecular Imaging.

**Closing Date: 20 February 2012**

**Faculty of Science and Engineering**
**Department of Chemistry**

## Postdoctoral Research Associate (Synthesis and Analysis of Peptide-Based Porous Materials)

**£31,020 - £35,939 pa**

Applications are invited for a Postdoctoral Research Associate in an experimental programme to develop porous materials based on peptides (Science 2010). You will have a PhD in Chemistry or Materials Science and expertise in inorganic and/or organic chemistry, synthesis of new materials and an excellent publication record. Experience in at least two of peptide chemistry, co-ordination chemistry, sorption measurements, solid-state NMR and diffraction methods are desirable. The post is available for 42 months (commencing March 2012).

**Job Ref: R-577010/N** **Closing date: 17 February 2012**

## Postdoctoral Research Associate (Computational Analysis of Peptide-Based Porous Materials)

**£31,020 - £35,939 pa**

Applications are invited for a Postdoctoral Research Associate in a computational programme to develop porous materials based on peptides. You will have a PhD in chemistry, physics or materials science, and excellent publication record and expertise in molecular dynamics/Monte Carlo and code development. Experience with DFT calculations and machine learning methods, are extremely desirable. The post is available for 36 months (commencing Sept 2012).

**Job Ref: R-577008/N** **Closing date: 17 February 2012**

**Faculty of Health and Life Sciences**
**Institute of Ageing and Chronic Disease**
**Department of Eye and Vision Science**

## Postdoctoral Research Associate

**£31,020 pa**

This 2 year project, funded by the Leverhulme Trust, is to investigate the control of saccadic eye movement in large groups of participants drawn from different cultural groups in order to establish whether recently reported differences in saccade control are due to cultural/environmental factors, or are culturally independent. The project involves recording and analysing saccades using various eyetrackers, and collation and analysis of large amounts of data. You should have a degree in Neuroscience, Psychology, Optometry, Orthoptics or a related discipline, with a relevant PhD. Previous experience of human visuomotor research, the use of visual stimulation, and particularly the use of eye tracking equipment, would be an advantage.

**Job Ref: R-577006/N** **Closing date: 17 February 2012**

U220732R

For full details, or to request an application pack, visit www.liv.ac.uk/working/job_vacancies/ or e-mail jobs@liv.ac.uk Please quote job ref in all enquiries.

COMMITTED TO DIVERSITY AND EQUALITY OF OPPORTUNITY

Stonewall DIVERSITY CHAMPION

POSITIVE ABOUT DISABLED PEOPLE

---

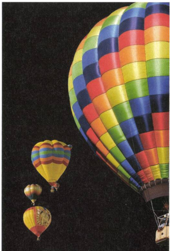
# INSTRUCTOR POSITION

The UAB Department of Pharmacology and Toxicology invites applications for a Faculty position at the Instructor level. Candidates with a PhD or equivalent degree and a strong record of research accomplishments will be considered. Applicants should submit a brief summary of research experience and interests, curriculum vitae and contact information for three professional references to Mary-Ann Bjornsti, Chair c/o Becky Warnix, to: Pharmtox.FacultyPositions @ccc.uab.edu.

*"UAB is an Equal Opportunity/Affirmative Action Employer committed to fostering a diverse, equitable and family-friendly environment in which all faculty and staff can excel and achieve work/life balance irrespective of ethnicity, gender, faith, and sexual orientation. UAB also encourages applications from individuals with disabilities and veterans."*

NW220483R

---



HIGHER & HIRE!

Post all your vacancies for FREE on the largest dedicated job board for the scientific community:

naturejobs.com

## Faculty Opportunities

UNIVERSITY OF CALIFORNIA
UC RIVERSIDE

### Assistant, Associate or Full Professor Positions in Biomedical Sciences

The Division of Biomedical Sciences within the new School of Medicine at the University of California Riverside is seeking to hire up to four new faculty members at the rank of Assistant, Associate or Full Professor. We are looking for accomplished research scientists in a number of specific areas including cancer, infectious disease, CNS neurological diseases, and cardiovascular and metabolic diseases. Preference will be given to those individuals examining molecular mechanisms of disease who are committed to a collaborative approach to research, and who are expert in their specific disease model.

The successful candidate will be appointed in the Division of Biomedical Sciences, joining a faculty who have directed a successful M.D. program in collaboration with UCLA, since the late 1970s. The new School of Medicine at UCR will be the sixth University of California Medical School and will serve the rapidly growing and dynamic Inland Southern California region by training a much-needed physician workforce and catalyzing innovations in research, education, and health care delivery that improve the health of medically underserved populations.

Areas of research within the Division include integrative immunology (vaccine development, neuro-immune, endocrine-immune, host- pathogen interactions), glial-neuronal interactions, neurodevelopmental disorders, cancer biology, cardiovascular disease, and diseases of ion transport. Particular strengths on the campus include genetics, epigenetics, genomics/bioinformatics, microRNAs, vector biology, bioengineering and nanotechnology, and synthetic and analytical chemistry.

The Division of Biomedical Sciences sponsors an innovative Ph.D. program that integrates the core medical curriculum with biomedical graduate training and research. The successful candidate will be expected to teach in the medical curriculum and actively participate in the Biomedical Sciences Ph.D. program. As such, preference will be given to candidates who are capable of teaching neuro- and/or general pharmacology, pathology, infectious disease (microbiology/virology), physiology (electrophysiology, renal or respiratory) or genetics.

The University of California, Riverside is situated in an historic citrus growing area surrounded by mountain ranges. Riverside is about an hour away from ski slopes, surfing, or hiking in mountain or desert environments, and housing in the area is very affordable. The campus is also located in a prime position to take advantage of the other universities, research institutes, and biotech industries present in Southern California.

Applicants must hold a Ph.D., M.D., Pharm D., or equivalent degree and qualify for a tenure-track/tenured faculty appointment at the University of California. Applications will be reviewed beginning February 27, 2012 and the positions will remain open until filled. UC Riverside is an Equal Opportunity/Affirmative Action employer.

### How to Apply

To apply, please submit the following items:

- Curriculum vitae
- Statement of research accomplishments and goals and teaching expertise
- Names of four individuals who will be asked to provide letters of reference once a short list is developed

Electronic submissions are encouraged. Send the items to:

**Violet Vargas**
Academic Personnel/HR Analyst
School of Medicine
University of California
Riverside, CA 92521
violet.vargas@ucr.edu

NW220613R

---

UCLA

## Computational Biosciences Initiative
## University of California Los Angeles
## Professor and Research Institute Director
## Position Announcement

UCLA's Computational Biosciences Institute is searching for a Director. The ideal candidate is a distinguished scholar who will bring intellectual leadership and synergy to this new UCLA Institute. The Institute is a Chancellor's Initiative, which is centered in the College of Letters and Sciences in partnership with the David Geffen School of Medicine (DGSOM), and participation from the Henry Samueli School of Engineering and Applied Science (HSSEAS), and the School of Public Health (SPH). This Initiative builds on more than 60 outstanding computational faculty members from more than 12 departments, many of who are affiliated with the recently established Interdepartmental PhD Program in Bioinformatics. The Initiative includes significant financial resources for new faculty appointments, renovated space for the new institute, and a collaboratory that offers bioinformatic resources. See website: www.cbi.ucla.edu.

The director's research is expected to lie within the areas of bioinformatics, computational biology, genomics, epigenomics, proteomics and evolutionary genomics, and should in part concern novel computational, quantitative, or bioinformatics methodology. The director must have an outstanding record of scholarly publications and research support. The successful candidate will hold a primary appointment in a department associated with Division of Life Sciences or Physical Sciences of the UCLA College, and the option of joint appointments in other academic units including the School of Medicine.

Letters of nomination and questions about the position should be sent to Dr. David Eisenberg at david.eisenberg@cbi.ucla.edu Application materials should be submitted on line through [www.mcdb.ucla.edu/compbiodir]. Please include a cover letter with a brief statement of research and vision, and CV. Review of applications will begin 10 February 2012. Please use position number 0865-1112-03 in all correspondence.

As a campus with a diverse student body, we encourage applications from women, minorities, and individuals with a history of mentoring under-represented minorities in the sciences. UCLA is an affirmative action/equal opportunity employer with a strong institutional commitment to the achievement of faculty and staff diversity.

NW220538R

---

Lupus
Research
Institute

## 2012 Distinguished Innovator Awards

The Lupus Research Institute invites applications for Distinguished Innovator Awards, a new global funding program that will run in parallel with our established Novel Research Grant program. Distinguished Innovator Awards will provide outstanding scientists with up to US $1 million to conduct novel research into the fundamental causes of lupus and lay the groundwork for a cure.

Scientists who have not previously worked in lupus are encouraged to apply, as are researchers working outside of the US.

Now offering two unique grant programs, the LRI is the world's leading private supporter of novel research in lupus.

Deadline for Letter of Intent: February 29, 2012
Deadline for Applications: April 2, 2012

For full guidelines visit our website
**www.lupusresearchinstitute.org**

NW220664A

**REPROGRAPHICS: CTT**, London E17 6BU UK and **The Charlesworth Group,** Wakefield, UK.
**PRINTED BY: William Gibbons & Sons Ltd,** UK; **Publishers Press**, Lebanon Junction, Ky, USA and **Obun Printing Co. Inc,** Tokyo, Japan

# A GAME OF SELF-DECEIT

*The resurrection shuffle*

**BY CLAYTON LOCKE**

I saw him duck under the cover of the awning, collapsing and shaking his umbrella, and step into the lazy smoke of the *izakaya*. When he raised his gaze to sweep the diners and drinkers at their tables, it took me a moment to work out what was wrong. Then I got it: I was looking at myself, not as I usually did, left–right inverted in the mirror, but as I saw myself in photographs. It's funny that my brain should pick out such a subtle detail. He saw me sitting at one of the outer tables, next to the plastic sheeting turned opaque with hammering rain. Then, as he moved away from the door, another figure entered. He was with my wife.

JACEY

His greeting bow was the slightest nod. "*Konbanwa*."

"Why did you bring her, you fool?" Without thinking I had addressed him using *temee*, an aggressive personal pronoun. I didn't even dare look at my wife, for fear of choking emotion.

"Because, Koji-*kun*, I want everything we do here to be witnessed. I'm not underestimating you."

I composed my face into the ever-necessary mask of neutrality. He had taken me off-guard, and I silently reprimanded myself. I harboured a sense of superiority over him, although I knew it to be illogical. This man was exactly me. Well, almost. I had the memories of the mountains and the time of cold, the experience that has shifted my entire way of thinking. A way of thinking I intended to preserve.

"This is illegal. You have to stay more than 20 kilometres away from me," he said.

I couldn't help but laugh at the notion that a bystander would pick us as clones; a salaryman in a crisply tailored suit, and a man in ill-fitting rags. He was right though, yet still he came. That's good, it meant he was afraid. I flicked a glance at my wife. For a heartbeat our eyes met and the blood sang in my ears.

"I've got a proposal for you," I said. I nodded to the empty chairs at my table and after a moment of hesitation they sat. I reached down for the plastic bag beside my chair and placed it on the table between us. Inside were documents and photos.

"So it's blackmail? I credited myself with more imagination than this. Where did you get these?" He paused a moment: when he spoke again his voice was level. "What's your problem? It's legal, you've been given a new identity, you're still *you*. You're wallowing in this self-serving misery, don't look to me to pull you out of it!"

"You don't know the half of it," I said. I reflected on all that had brought me here, how I should never have gone mountaineering without a backup hoverbelt. It was only when the primary one didn't catch and I slipped and fell that I'd realized how stupid I'd been. I'd come to a stop in a deep ravine, my knees shattered, coughing blood. A blow to my head had cracked the personal smartphone chip embedded in my skull. Suddenly I was cut off from the world, no Internet access, no emergency calls, and a silence like I'd not known in many years. I'd been missing for less than two days, but who could survive those subzero temperatures? Lost in the mist and rain, hearing the electric hiss of hover-copters as they glided overhead, and shouting myself hoarse. I'd finally been found, but by the time the gears of bureaucracy had turned, nearly two weeks had passed, and to the world I was dead.

**➲ NATURE.COM**
Follow Futures on Facebook at:
**go.nature.com/mtoodm**

How easily they'd discarded me. Life was cheap now, but death was cheaper. Already a clone body had been grown and my latest memory backup imprinted on the fresh brain. It wasn't possible to re-enter my life, for my new self had all the rights and identity: everything from bank accounts and job qualifications to friends and my wife.

I had changed in those weeks. Although I cursed myself for my stupidity, I held the experience in the mountains as the most important thing that had ever happened to me. I wanted to keep it and shape my life with it as part of me. I'd spent long nights seething with jealousy at the imposter taking over my life. Then I had a plan, and the very next day I spent the last yen of my pension on a full up-to-the-minute memory backup at a professional storage agency.

I blinked back into the present. The new me snatched across the table and tried to grab the plastic bag but I caught his wrist and in a moment we were both on our feet and struggling. In the confusion I pulled the gun from my rear pocket. This is where my preparations paid off, my risky burglary of what had once been my own home to retrieve this very weapon.

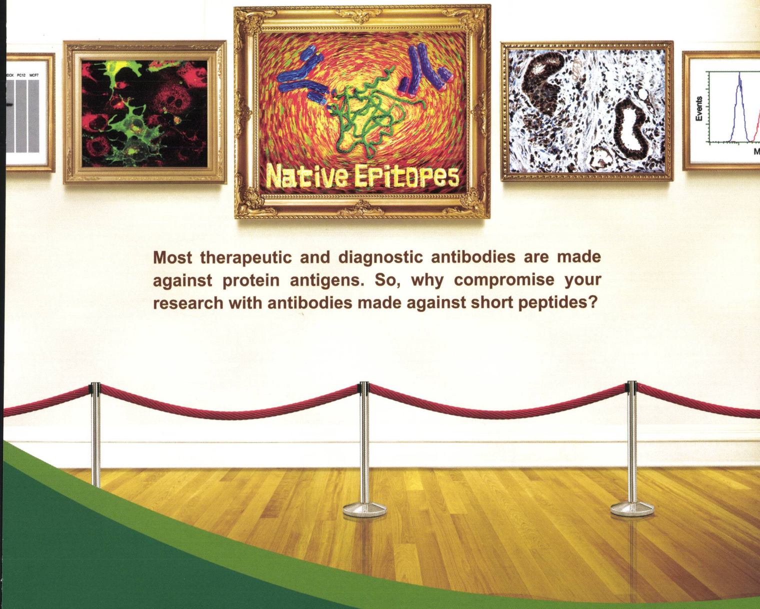
The gun was between us now (it was an evenly matched fight, I observed wryly) and I sensed a crowd gathering. I twisted backwards, keeping my hold on him, and we fell to the ground, and the gun was pointed at my own temple, my finger around the trigger. I didn't think I'd have the courage, but now in the moment it was all too easy.

"Murderers don't come back," I said. I saw his eyes widen as he recognized his pistol, the realization he'd been out-played. He knew what I knew, the death penalty was sure and a criminal would not be resurrected. It would be the *real* me, who had lived through the mountain cold, who would live again. He tried to withdraw his hands from the pistol but it was too late. The muzzle tight against the side of my head, I pulled the trigger. ■

**Clayton Locke** *is an Australian physicist currently researching frequency standards and metrology in Tokyo.*

# TrueMAB™ Monoclonal Antibodies

## made against authentic protein antigens



**Most therapeutic and diagnostic antibodies are made against protein antigens. So, why compromise your research with antibodies made against short peptides?**

## TrueMAB™ recognizing protein native epitopes

- **Releasing over 200 TrueMAB™ monthly**
- **Extensive validation ( IHC on 24 FFPE tissues, IF, Flow, IP, WB )**
- **Positive controls included**

ORIGENE
Your Gene Company
origene.com/truemab

applied biosystems
by life technologies™
reliable service & support
intuitive protocol setup
precise temperature control
color touch screen
six annealing temperatures
more control
at your fingertips
Applied Biosystems
Veriti

## Easy PCR optimization is just a touch away

The Veriti® Thermal Cycler features powerful innovations for precise control of your PCR optimization, including a color touch screen and intuitive interface, and VeriFlex™ Blocks—six independent temperature blocks—for more flexibility and even greater precision. All with the reliability you expect from Life Technologies.

Get more control of your PCR optimization at
**www.appliedbiosystems.com/veriti**
**Get the free mobile app for your phone at http://gettag.mobi**

life technologies™

©2011 Life Technologies Corporation. All rights reserved. The trademarks mentioned herein are the property of Life Technologies Corporation or their respective owners. For research use only. Not intended for any animal or human therapeutic or diagnostic use. CO31272 1011